OM of: US-09-303-518D-569 to: SPTREMBL_19:* out_format : pfs

Date: Jun 30, 2002 8:13 AM

About: Results were produced by the GenCore software, version 4.5,
Copyright (c) 1993-2000 Compugen Ltd.

Command line parameters:
-MODEL=frame+_n2p.model -DEV=x1h
-Q=/cgn2_1/USPTO_spool/US09303518/runat_28062002_142714_4338/app_query.fasta_1.23501
-DB=SPTREMBL_19 -QFMT=fastan -SUFFIX=rspt -GAPOP=12.000
-GAPEXT=4.000 -MINMATCH=0.100 -LOOPCL=0.000 -LOOPEXT=0.000
-QGAPOP=4.500 -QGAPEXT=0.050 -XGAPOP=10.000 -XGAPEXT=0.500
-FGAPOP=6.000 -FGAPEXT=7.000 -YGAPOP=10.000 -YGAPEXT=0.500
-DELOP=6.000 -DELEXT=7.000 -START=1 -MATRIX=blosum62
-TRANS=human40.cdi -LIST=100 -DOCALIGN=200 -THR_SCORE=pct
-THR_MAX=100 -THR_MIN=0 -ALIGN=45 -MODE=LOCAL -OUTFMT=pfs
-NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09303518_@CGN1_1_1967 -NCPU=6 -ICPU=3 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -NO_XLPXY -WAIT -THREADS=1

Search information block:
Query: US-09-303-518D-569
Query length: 897
Database: SPTREMBL_19:*
Database sequences: 562222
Database length: 172994929
Search time (sec): 883.700000

score_list:

Sequence	Strd	Orig	ZScore	EScore	Len	Documentation
sp_bacteriap:Q9JVV7	+	1580.00	2512.91	4.3e-132	298	! Q9jvv7 neisseria meningitidis (
sp_bacteriap:Q9JY07	+	1577.00	2508.12	7.9e-132	298	! Q9jy07 neisseria meningitidis (
sp_bacteria:Q9EYM0	+	293.00	455.70	1.5e-17	337	! Q9eym0 acinetobacter sp. bd413
sp_bacteriap:Q9PDN1	+	263.50	409.20	6.0e-15	316	! Q9pdn1 xylella fastidiosa. lipi
sp_bacteriap:Q9I7B5	+	259.00	402.68	1.5e-14	295	! Q9i7b5 pseudomonas aeruginosa.
sp_bacteriap:Q9HYZ8	+	228.50	353.41	7.8e-12	312	! Q9hyz8 pseudomonas aeruginosa.
sp_bacteria:O33437	+	221.50	344.28	3.1e-11	253	! O33437 pseudomonas fluorescens.
sp_bacteriap:Q9JYV2	+	204.50	315.82	1.0e-09	289	! Q9jyv2 neisseria meningitidis (
sp_bacteriap:Q9JTU4	+	203.50	314.22	1.3e-09	289	! Q9jtu4 neisseria meningitidis (
sp_bacteriap:Q9PH42	+	196.50	303.38	5.4e-09	279	! Q9ph42 xylella fastidiosa. lipi
sp_bacteria:Q9API8	+	172.50	264.60	7.4e-07	292	! Q9api8 methylobacterium extorqu
sp_bacteria:Q9L7U9	+	172.50	263.76	7.6e-07	318	! Q9l7u9 haemophilus ducreyi. lip
sp_bacteriap:O84013	+	172.50	260.26	8.3e-07	455	! O84013 chlamydia trachomatis. a
sp_bacteriap:Q9ZCL1	+	169.50	259.87	1.4e-06	290	! Q9zcl1 rickettsia prowazekii. l
sp_bacteria:Q9AKA6	+	166.00	254.28	2.8e-06	290	! Q9aka6 rickettsia typhi. htrb p
sp_bacteriap:Q9PL30	+	164.50	247.52	4.3e-06	453	! Q9pl30 chlamydia muridarum. lip
sp_bacteriap:Q9CME5	+	163.50	249.54	4.8e-06	313	! Q9cme5 pasteurella multocida. H
sp_bacteriap:Q9CNR4	+	162.50	247.79	5.9e-06	318	! Q9cnr4 pasteurella multocida. m
sp_bacteria:Q9L7Q3	+	161.50	246.22	7.2e-06	317	! Q9l7q3 haemophilus ducreyi. htr
sp_bacteria:Q9AKF0	+	154.00	235.11	3.3e-05	290	! Q9akf0 rickettsia rickettsii. H
sp_bacteriap:Q98L49	+	153.00	232.31	4.1e-05	328	! Q98l49 rhizobium loti (mesorhiz
sp_bacteria:O82937	+	151.00	228.99	6.3e-05	332	! O82937 escherichia coli o157:h7
sp_bacteria:Q9ZGU3	+	151.00	228.68	6.3e-05	343	! Q9zgu3 escherichia coli o157:h7
sp_bacteriap:Q92GN2	+	149.00	227.12	9.1e-05	290	! Q92gn2 rickettsia conorii. lipi
sp_bacteria:P94909	+	147.50	218.84	0.0001	529	! P94909 microbacterium ammoniaph
sp_bacteria:Q9AKK5	+	147.00	223.93	0.0001	290	! Q9akk5 rickettsia montana. htrb
sp_bacteriap:Q9K217	+	146.00	217.77	0.0002	462	! Q9k217 chlamydia pneumoniae (ch
sp_bacteriap:Q9Z983	+	146.00	217.67	0.0002	467	! Q9z983 chlamydia pneumoniae (ch
sp_bacteriap:Q9KVD4	+	141.00	213.23	0.0005	325	! Q9kvd4 vibrio cholerae. lipid a
sp_bacteriap:Q9KVD3	+	136.50	206.25	0.0012	318	! Q9kvd3 vibrio cholerae. lipid a
sp_bacteriap:Q92P56	+	135.50	204.84	0.0015	312	! Q92p56 rhizobium meliloti (sinc
sp_bacteria:Q9F6E3	+	135.00	203.79	0.0016	320	! Q9f6e3 streptomyces collinus. p
sp_bacteria:O06470	+	133.50	199.09	0.0024	405	! O06470 burkholderia cepacia (ps
sp_bacteria:Q51415	+	129.50	187.57	0.0062	684	! Q51415 pseudomonas aeruginosa.
sp_human:Q9UQ39	+	128.50	182.70	0.0082	956	! Q9uq39 homo sapiens (human). rn
sp_human:Q9UQ40	+	128.50	179.98	0.0088	1262	! Q9uq40 homo sapiens (human). r
sp_human:Q9UHA8	+	128.50	174.12	0.0103	2296	! Q9uha8 homo sapiens (human). s
sp_human:Q9UQ35	+	128.50	172.35	0.0108	2752	! Q9uq35 homo sapiens (human). r
sp_plant:Q9ARV4	+	128.00	190.55	0.0073	395	! Q9arv4 oryza sativa (rice). p05
sp_bacteria:Q9ZI82	+	127.50	191.72	0.0077	323	! Q9zi82 salmonella typhimurium.
sp_human:Q16824	+	125.50	179.69	0.0146	797	! Q16824 homo sapiens (human).
sp_bacteria:Q52747	+	125.00	180.44	0.0155	680	! Q52747 ruminococcus flavefac
sp_bacteria:Q53478	+	124.00	182.75	0.0172	456	! Q53478 neisseria gonorrhoeae
sp_bacteria:Q9LCV5	+	123.50	175.45	0.0225	886	! Q9lcv5 rhodospirillum rubrum
sp_bacteria:Q9RA26	+	122.00	183.24	0.0235	313	! Q9ra26 vibrio marinus (morit
sp_human:Q9NR59	+	120.50	166.82	0.0460	1312	! Q9nr59 homo sapiens (human)
sp_human:Q9H7N4	+	120.50	166.59	0.0463	1343	! Q9h7n4 homo sapiens (human)
sp_virus:Q9WFA2	+	119.50	170.22	0.0496	787	! Q9wfa2 snow goose hepatitis
sp_virus:Q9WFA5	+	119.50	170.22	0.0496	787	! Q9wfa5 snow goose hepatitis
sp_virus:Q9WFA8	+	119.50	170.22	0.0496	787	! Q9wfa8 snow goose hepatitis
sp_virus:Q9WFB2	+	119.50	170.22	0.0496	787	! Q9wfb2 snow goose hepatitis
sp_virus:Q9WFB5	+	119.50	170.22	0.0496	787	! Q9wfb5 snow goose hepatitis
sp_bacteriap:Q98DE3	+	119.00	171.36	0.0522	646	! Q98de3 rhizobium loti (mesor
sp_human:Q9H9J4	+	117.50	162.92	0.0832	1198	! Q9h9j4 homo sapiens (human)
sp_human:Q9H4U3	+	116.50	165.30	0.0920	798	! Q9h4u3 homo sapiens (human).
sp_bacteriap:Q9RYP2	+	116.00	171.86	0.0841	376	! Q9ryp2 deinococcus radiodura
sp_bacteriap:O07148	+	115.50	172.64	0.0894	320	! O07148 mycobacterium leprae.
sp_bacteria:Q9AFL5	+	115.00	172.03	0.0986	314	! Q9afl5 shigella flexneri. ac
sp_bacteria:Q9AJU9	+	115.00	172.00	0.0987	315	! Q9aju9 shigella flexneri. ms
sp_bacteria:O05765	+	114.50	164.84	0.1290	603	! O05765 mycobacterium smegmat
sp_bacteriap:Q9A645	+	114.00	170.43	0.1210	314	! Q9a645 caulobacter crescentu
sp_bacteria:Q9RFF2	+	114.00	164.61	0.1409	569	! Q9rff2 rhodobacter sphaeroid
sp_rodent:Q99053	+	114.00	163.69	0.1443	625	! Q99053 rattus norvegicus (ra
sp_human:Q9BXI1	+	114.00	158.86	0.1637	1024	! Q9bxi1 homo sapiens (human)
sp_human:Q96EL5	+	114.00	157.98	0.1675	1120	! Q96el5 homo sapiens (human)
sp_invertebrate:Q9VC95	+	113.50	164.47	0.1534	532	! Q9vc95 drosophila melanoga
sp_bacteria:O68391	+	113.50	164.29	0.1541	542	! O68391 pseudomonas pseudoalc
sp_bacteria:Q9L283	+	113.00	168.93	0.1482	311	! Q9l283 streptomyces coelicol
sp_bacteriap:Q9A3Z5	+	113.00	160.47	0.1848	738	! Q9a3z5 caulobacter crescentu
sp_rodent:O70162	+	113.00	159.51	0.1895	814	! O70162 mus musculus (mouse).
sp_invertebrate:Q9W3P3	+	113.00	152.34	0.2284	1693	! Q9w3p3 drosophila melanog
sp_rodent:O88778	+	113.00	144.08	0.2833	3938	! O88778 rattus norvegicus (r
sp_invertebrate:Q9VU43	+	112.50	157.16	0.2186	954	! Q9vu43 drosophila melanoga
sp_plant:Q9SM80	+	112.50	155.81	0.2264	1095	! Q9sm80 oryza sativa (rice).
sp_plant:Q943U4	+	112.50	152.85	0.2445	1481	! Q943u4 oryza sativa (rice).
sp_bacteria:O85018	+	112.00	154.77	0.2524	1122	! O85018 mycobacterium smegma
sp_human:Q9UQ36	+	112.00	153.52	0.2608	1275	! Q9uq36 homo sapiens (human)
sp_human:O15038	+	112.00	150.24	0.2841	1783	! O15038 homo sapiens (human)
sp_human:O60382	+	112.00	150.19	0.2844	1791	! O60382 homo sapiens (human)
sp_bacteriap:Q9ZMF5	+	111.00	165.21	0.2263	328	! Q9zmf5 helicobacter pylori j
sp_plant:Q94J70	+	110.50	166.54	0.2372	264	! Q94j70 oryza sativa (rice).
sp_invertebrate:Q9NE62	+	110.50	148.21	0.3826	1718	! Q9ne62 leishmania major.
sp_plant:Q94DC9	+	109.00	160.58	0.3540	380	! Q94dc9 oryza sativa (rice).
sp_invertebrate:Q9NC84	+	109.00	160.15	0.3579	397	! Q9nc84 strongylocentrotus
sp_bacteriap:Q98J30	+	109.00	158.13	0.3773	488	! Q98j30 rhizobium loti (mesor
sp_bacteriap:Q9HZM8	+	109.00	150.56	0.4596	1057	! Q9hzm8 pseudomonas aerugino
sp_plant:Q942Z9	+	108.50	155.71	0.4361	576	! Q942z9 oryza sativa (rice).
sp_bacteriap:Q9HUI3	+	108.00	149.53	0.5559	998	! Q9hui3 pseudomonas aeruginos
sp_plant:Q94HD1	+	107.50	158.05	0.4829	385	! Q94hd1 oryza sativa (rice).
sp_rodent:Q9Z2N3	+	107.50	153.52	0.5436	612	! Q9z2n3 mus musculus (mouse).
sp_plant:Q947V1	+	107.50	151.95	0.5662	718	! Q947v1 oryza sativa (rice).
sp_plant:Q9SKR5	+	107.50	151.55	0.5722	748	! Q9skr5 arabidopsis thaliana
sp_plant:Q39177	+	107.50	151.32	0.5757	766	! Q39177 arabidopsis thaliana
sp_human:O75370	+	107.00	159.00	0.5112	322	! O75370 homo sapiens (human).
sp_bacteriap:O25057	+	107.00	158.82	0.5136	328	! O25057 helicobacter pylori (
sp_bacteriap:Q9I1T4	-	107.00	157.46	0.5322	377	! Q9i1t4 pseudomonas aeruginos
sp_human:O60585	+	107.00	149.85	0.6489	820	! O60585 homo sapiens (human).
sp_plant:Q94JG5	+	106.50	152.96	0.6493	550	! Q94jg5 oryza sativa (rice).
sp_vertebrate:Q90783	+	106.50	145.55	0.7878	1173	! Q90783 gallus gallus (chick
sp_bacteria:Q93SC4	+	106.00	162.72	0.5461	187	! Q93sc4 frankia sp. ari3. hyp

seq_name: sp_bacteriap:Q9JVV7

seq_documentation_block:
ID Q9JVV7 PRELIMINARY; PRT; 298 AA.
AC Q9JVV7;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE PUTATIVE ACYLTRANSFERASE.
GN NMA0662.
OS Neisseria meningitidis (serogroup A).
OC Bacteria; Proteobacteria; beta subdivision; Neisseriaceae; Neisseria.

```
OX   NCBI_TaxID=65699;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Z2491 / SEROGROUP A / SEROTYPE 4A;
RX   MEDLINE=20222556; PubMed=10761919;
RA   Parkhill J., Achtman M., James K.D., Bentley S.D., Churcher C.,
RA   Klee S.R., Morelli G., Basham D., Brown D., Chillingworth T.,
RA   Davies R.M., Davis P., Devlin K., Feltwell T., Hamlin N., Holroyd S.,
RA   Jagels K., Leather S., Moule S., Mungall K., Quail M.A.,
RA   Rajandream M.A., Rutherford K.M., Simmonds M., Skelton J.,
RA   Whitehead S., Spratt B.G., Barrell B.G.;
RT   "Complete DNA sequence of a serogroup A strain of Neisseria
RT   meningitidis Z2491.";
RL   Nature 404:502-506(2000).
DR   EMBL; AL162753; CAB83949.1; -.
KW   Transferase; Acyltransferase; Complete proteome.
SQ   SEQUENCE   298 AA;  33510 MW;  2A09946670E283C2 CRC64;


alignment_scores:
            Quality: 1580.00              Length:    298
              Ratio:  5.302                Gaps:      0
 Percent Similarity: 100.000   Percent Identity: 100.000

alignment_block:
 US-09-303-518D-569 x Q9JVV7   ..

 Align seg 1/1  to: Q9JVV7  from: 1  to: 298

       1 ATGTTTCGTTTACAATTCAGGCTGTTTCCCCCTTTGCGAACCGCCATGCA 50
         ||||||||||||||||||||||||||||||||||||||||||||||||||
       1 MetPheArgLeuGlnPheArgLeuPheProProLeuArgThrAlaMetHi 17

      51 CATCCTGTTGACCGCCCTGCTCAAATGCCTCTCCCTGCTGCCGCTTTCCT 100
         ||||||||||||||||||||||||||||||||||||||||||||||||||
      17 sIleLeuLeuThrAlaLeuLeuLysCysLeuSerLeuLeuProLeuSerC 34

     101 GTCTGCACACGCTGGGAAACCGGCTCGGACATCTGGCGTTTTACCTTTTA 150
         ||||||||||||||||||||||||||||||||||||||||||||||||||
      34 ysLeuHisThrLeuGlyAsnArgLeuGlyHisLeuAlaPheTyrLeuLeu 50

     151 AAGGAAGACCGCGCGCGCATCGTCGCCAATATGCGTCAGGCAGGCATGAA 200
         ||||||||||||||||||||||||||||||||||||||||||||||||||
      51 LysGluAspArgAlaArgIleValAlaAsnMetArgGlnAlaGlyMetAs 67

     201 TCCCGACCCCAAAACGGTCAAAGCCGTTTTTGCGGAAACGGCAAAAGGCG 250
         ||||||||||||||||||||||||||||||||||||||||||||||||||
      67 nProAspProLysThrValLysAlaValPheAlaGluThrAlaLysGlyG 84

     251 GTTTGGAACTTGCCCCCGCGTTTTTCAGAAAACCGGAAGACATAGAAACA 300
         ||||||||||||||||||||||||||||||||||||||||||||||||||
      84 lyLeuGluLeuAlaProAlaPhePheArgLysProGluAspIleGluThr 100

     301 ATGTTCAAAGCGGTACACGGCTGGGAACATGTGCAGCAGGCTTTGGACAA 350
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     101 MetPheLysAlaValHisGlyTrpGluHisValGlnGlnAlaLeuAspLy 117

     351 ACACGAAGGGCTGCTATTCATCACGCCGCACATCGGCAGCTACGATTTGG 400
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     117 sHisGluGlyLeuLeuPheIleThrProHisIleGlySerTyrAspLeuG 134

     401 GCGGACGCTACATCAGCCAGCAGCTTCCGTTCCCGCTGACCGCCATGTAC 450
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     134 lyGlyArgTyrIleSerGlnGlnLeuProPheProLeuThrAlaMetTyr 150

     451 AAACCGCCGAAAATCAAAGCGATAGACAAAATCATGCAGGCGGGCAGGGT 500
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     151 LysProProLysIleLysAlaIleAspLysIleMetGlnAlaGlyArgVa 167

     501 TCGCGGCAAAGGAAAAACCGCGCCTACCAGCATACAAGGGGTCAAACAAA 550
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     167 lArgGlyLysGlyLysThrAlaProThrSerIleGlnGlyValLysGlnI 184

     551 TCATCAAAGCCCTGCGTTCGGGCGAAGCAACCATCGTCCTGCCCGACCAC 600
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     184 leIleLysAlaLeuArgSerGlyGluAlaThrIleValLeuProAspHis 200

     601 GTCCCCTCCCCTCAAGAAGGCGGGGAAGGCGTATGGGTGGATTTCTTCGG 650
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     201 ValProSerProGlnGluGlyGlyGluGlyValTrpValAspPhePheGl 217

     651 CAAACCTGCCTATACCATGACGCTGGCGGCAAAATTGGCACACGTCAAAG 700
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     217 yLysProAlaTyrThrMetThrLeuAlaAlaLysLeuAlaHisValLysG 234

     701 GCGTGAAAACCCTGTTTTTCTGCTGCGAACGCCTGCCTGGCGGACAAGGT 750
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     234 lyValLysThrLeuPhePheCysCysGluArgLeuProGlyGlyGlnGly 250

     751 TTCGATTTGCACATCCGCCCCGTCCAAGGGGAATTGAACGGCGACAAAGC 800
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     251 PheAspLeuHisIleArgProValGlnGlyGluLeuAsnGlyAspLysAl 267

     801 CCATGATGCCGCCGTGTTCAACCGCAATGCCGAATATTGGATACGCCGTT 850
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     267 aHisAspAlaAlaValPheAsnArgAsnAlaGluTyrTrpIleArgArgP 284

     851 TTCCGACGCAGTATCTGTTTATGTACAACCGCTACAAAATGCCG 894
         |||||||||||||||||||||||||||||||||||||||||||
     284 heProThrGlnTyrLeuPheMetTyrAsnArgTyrLysMetPro 298

seq_name: sp_bacteriap:Q9JY07

seq_documentation_block:
ID   Q9JY07      PRELIMINARY;      PRT;   298 AA.
AC   Q9JY07;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   HTRB/MSBB FAMILY PROTEIN.
GN   NMB1801.
OS   Neisseria meningitidis (serogroup B).
OC   Bacteria; Proteobacteria; beta subdivision; Neisseriaceae; Neisseria.
OX   NCBI_TaxID=491;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=MC58 / SEROGROUP B;
RX   MEDLINE=20175755; PubMed=10710307;
RA   Tettelin H., Saunders N.J., Heidelberg J., Jeffries A.C., Nelson K.E.,
RA   Eisen J.A., Ketchum K.A., Hood D.W., Peden J.F., Dodson R.J.,
RA   Nelson W.C., Gwinn M.L., DeBoy R., Peterson J.D., Hickey E.K.,
RA   Haft D.H., Salzberg S.L., White O., Fleischmann R.D., Dougherty B.A.,
RA   Mason T., Ciecko A., Parksey D.S., Blair E., Cittone H., Clark E.B.,
RA   Cotton M.D., Utterback T.R., Khouri H., Qin H., Vamathevan J.,
RA   Gill J., Scarlato V., Masignani V., Pizza M., Grandi G., Sun L.,
RA   Smith H.O., Fraser C.M., Moxon E.R., Rappuoli R., Venter J.C.;
RT   "Complete genome sequence of Neisseria meningitidis serogroup B strain
RT   MC58.";
RL   Science 287:1809-1815(2000).
DR   EMBL; AE002530; AAF42138.1; -.
DR   TIGR; NMB1801; -.
KW   Complete proteome.
SQ   SEQUENCE   298 AA;  33492 MW;  013437BC4DD6508F CRC64;


alignment_scores:
            Quality: 1577.00              Length:    298
              Ratio:  5.292                Gaps:      0
 Percent Similarity: 100.000   Percent Identity: 99.664

alignment_block:
 US-09-303-518D-569 x Q9JY07   ..

 Align seg 1/1  to: Q9JY07  from: 1  to: 298

       1 ATGTTTCGTTTACAATTCAGGCTGTTTCCCCCTTTGCGAACCGCCATGCA 50
```

```
      ||||||||||||||||||||||||||||||||||||||||||||||||||
    1 MetPheArgLeuGlnPheArgLeuPheProProLeuArgThrAlaMetHi 17

   51 CATCCTGTTGACCGCCCTGCTCAAATGCCTCTCCCTGCTGCCGCTTTCCT 100
      ||||||||||||||||||||||||||||||||||||||||||||||||||
   17 sIleLeuLeuThrAlaLeuLeuLysCysLeuSerLeuLeuProLeuSerC 34

  101 GTCTGCACACGCTGGGAAACCGGCTCGGACATCTGGCGTTTTACCTTTTA 150
      ||||||||||||||||||||||||||||||||||||||||||||||||||
   34 ysLeuHisThrLeuGlyAsnArgLeuGlyHisLeuAlaPheTyrLeuLeu 50

  151 AAGGAAGACCGCGCGCGCATCGTCGCCAATATGCGTCAGGCAGGCATGAA 200
      ||||||||||||||||||||||||||||||||||||||||||||||:::||
   51 LysGluAspArgAlaArgIleValAlaAsnMetArgGlnAlaGlyLeuAs 67

  201 TCCCGACCCCAAAACGGTCAAAGCCGTTTTTGCGGAAACGGCAAAAGGCG 250
      ||||||||||||||||||||||||||||||||||||||||||||||||||
   67 nProAspProLysThrValLysAlaValPheAlaGluThrAlaLysGlyG 84

  251 GTTTGGAACTTGCCCCCGCGTTTTTCAGAAAACCGGAAGACATAGAAACA 300
      ||||||||||||||||||||||||||||||||||||||||||||||||||
   84 lyLeuGluLeuAlaProAlaPhePheArgLysProGluAspIleGluThr 100

  301 ATGTTCAAAGCGGTACACGGCTGGGAACATGTGCAGCAGGCTTTGGACAA 350
      ||||||||||||||||||||||||||||||||||||||||||||||||||
  101 MetPheLysAlaValHisGlyTrpGluHisValGlnGlnAlaLeuAspLy 117

  351 ACACGAAGGGCTGCTATTCATCACGCCGCACATCGGCAGCTACGATTTGG 400
      ||||||||||||||||||||||||||||||||||||||||||||||||||
  117 sHisGluGlyLeuLeuPheIleThrProHisIleGlySerTyrAspLeuG 134

  401 GCGGACGCTACATCAGCCAGCAGCTTCCGTTCCCGCTGACCGCCATGTAC 450
      ||||||||||||||||||||||||||||||||||||||||||||||||||
  134 lyGlyArgTyrIleSerGlnGlnLeuProPheProLeuThrAlaMetTyr 150

  451 AAACCGCCGAAAATCAAAGCGATAGACAAAATCATGCAGGCGGGCAGGGT 500
      ||||||||||||||||||||||||||||||||||||||||||||||||||
  151 LysProProLysIleLysAlaIleAspLysIleMetGlnAlaGlyArgVa 167

  501 TCGCGGCAAAGGAAAAACCGCGCCTACCAGCATACAAGGGGTCAAACAAA 550
      ||||||||||||||||||||||||||||||||||||||||||||||||||
  167 lArgGlyLysGlyLysThrAlaProThrSerIleGlnGlyValLysGlnI 184

  551 TCATCAAAGCCCTGCGTTCGGGCGAAGCAACCATCGTCCTGCCCGACCAC 600
      ||||||||||||||||||||||||||||||||||||||||||||||||||
  184 leIleLysAlaLeuArgSerGlyGluAlaThrIleValLeuProAspHis 200

  601 GTCCCCTCCCCTCAAGAAGGCGGGGAAGGCGTATGGGTGGATTTCTTCGG 650
      ||||||||||||||||||||||||||||||||||||||||||||||||||
  201 ValProSerProGlnGluGlyGlyGluGlyValTrpValAspPhePheGl 217

  651 CAAACCTGCCTATACCATGACGCTGGCGGCAAAATTGGCACACGTCAAAG 700
      ||||||||||||||||||||||||||||||||||||||||||||||||||
  217 yLysProAlaTyrThrMetThrLeuAlaAlaLysLeuAlaHisValLysG 234

  701 GCGTGAAAACCCTGTTTTTCTGCTGCGAACGCCTGCCTGGCGGACAAGGT 750
      ||||||||||||||||||||||||||||||||||||||||||||||||||
  234 lyValLysThrLeuPhePheCysCysGluArgLeuProGlyGlyGlnGly 250

  751 TTCGATTTGCACATCCGCCCCGTCCAAGGGGAATTGAACGGCGACAAAGC 800
      ||||||||||||||||||||||||||||||||||||||||||||||||||
  251 PheAspLeuHisIleArgProValGlnGlyGluLeuAsnGlyAspLysAl 267

  801 CCATGATGCCGCCGTGTTCAACCGCAATGCCGAATATTGGATACGCCGTT 850
      ||||||||||||||||||||||||||||||||||||||||||||||||||
  267 aHisAspAlaAlaValPheAsnArgAsnAlaGluTyrTrpIleArgArgP 284

  851 TTCCGACGCAGTATCTGTTTATGTACAACCGCTACAAAATGCCG 894
      |||||||||||||||||||||||||||||||||||||||||||
  284 heProThrGlnTyrLeuPheMetTyrAsnArgTyrLysMetPro 298

seq_name: sp_bacteria:Q9EYM0
```

```
seq_documentation_block:
ID   Q9EYM0      PRELIMINARY;      PRT;   337 AA.
AC   Q9EYM0;
DT   01-MAR-2001 (TrEMBLrel. 16, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   PUTATIVE HTRB PROTEIN.
OS   Acinetobacter sp. BD413.
OC   Bacteria; Proteobacteria; gamma subdivision; Moraxellaceae;
OC   Acinetobacter.
OX   NCBI_TaxID=104611;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=BD413;
RX   MEDLINE=21318707; PubMed=11425734;
RA   Friedrich A., Hartsch T., Averhoff B.;
RT   "Natural' Transformation in Mesophilic and Thermophilic Bacteria:
RT   Identification and Characterization of Novel, Closely Related
RT   Competence Genes in Acinetobacter sp. Strain BD413 and Thermus
RT   thermophilus HB27.";
RL   Appl. Environ. Microbiol. 67:3140-3148(2001).
DR   EMBL; AF320001; AAG34711.1; -.
DR   InterPro; IPR000515; BPD_transp.
DR   PROSITE; PS00402; BPD_TRANSP_INN_MEMBR; UNKNOWN_1.
SQ   SEQUENCE   337 AA;  38795 MW;  6E998AFED1CEB140 CRC64;


alignment_scores:
            Quality:  293.00              Length:    193
              Ratio:  2.170                 Gaps:      5
 Percent Similarity:  69.948   Percent Identity: 35.233

alignment_block:
 US-09-303-518D-569 x Q9EYM0   ..

 Align seg 1/1  to: Q9EYM0  from: 1  to: 337

  313 GTACACGGCTGGGAACATGTGCAGCAGGCTTTGGACAAACACGAAGGGCT 362
      |||:::|||   :::      :::||||||||:::  ::::::::::|||::
  100 ValGluGlyAlaAspPhePheHisGlnAlaIleAlaGluAsnLysGlyIl 116

  363 GCTATTCATCACGCCGCACATCGGCAGCTACGATTTGGGCGGACGCTACA 412
      :::::::|||:::|||||||:::|||:::::::::::::   :::   :::
  116 eValLeuIleValProHisPheGlyThrTrpGluIleMetAsnAlaTrpC 133

  413 TCAGCCAGCAGCTTCCGTTCCCGCTGACCGCCATGTACAAACCGCCGAAA 462
        :::|||               :::|||   :::||||||||||   |||
  133 ysAlaGlnTyrThr......AspMetThrIleLeuTyrLysProValLys 147

  463 ATCAAAGCGATAGACAAAATCATGCAGGCGGGCAGGGTTCGCGGCAAAGG 512
        |||      |||:::::::::::::   :::|||   |||   :::::
  148 AspLysAspAlaAspArgPheValArgGluAlaArgSerArgGluGlnAl 164

  513 AAAAACCGCGCCTACCAGCATACAAGGGGTCAAACAAATCATCAAAGCCC 562
      :|||   :::|||||||:::   :::||||||:::|||||||:::|||||||
  164 aLysLeuValProThrAspGluSerGlyValArgGlnIlePheLysAlaL 181

  563 TGCGTTCGGGCGAAGCAACCATCGTCCTGCCCGACCACGTCCCCTCCCCT 612
      ||:::::|||   :::|||:::::|||||||||||||      :::|||
  181 euLysGlnGlyGlyThrThrValIleLeuProAspHis......ThrPro 195

  613 CAAGAAGGCGGGGAAGGCGTATGGGTGGATTTCTTCGGCAAACCTGCCTA 662
      :::   ||||||||         ||||||:::||||||   |||
  196 AsnValGlyGlyGlu......MetValAspTyrPheGlyIleProLeuAl 210

  663 TACCATGACGCTGGCGGCAAAATTGGCACACGTCAAAGGCGTGAAAACCC 712
       :::   :::|||:::|||||||||   :::         :::|||:::|
  210 aSerSerAsnLeuSerAlaLysLeuIleGlnLysThrLysAlaLysAlaL 227

  713 TGTTTTTCTGCTGCGAACGCCTGCCTGGCGGACAAGGTTTCGATTTGCAC 762
      ||||||:::   :::   |||         ::::::||||||||:::
  227 euPheLeuTyrAlaIleArg...AsnGluAsnHisGlyPheAspMetPhe 242
```

```
    763 ATCCGCCCCGTCCAAGGGGAATTGAACGGCGACAAAGCCCATGATGCCGC 812
        |||:::|||::::::      :::                |||||||:::::
    243 IleGluProMetAspProAlaIleTyrGlnGlyThrAspHisAspGlyTh 259

    813 C...GTGTTCAACCGCAATGCCGAATATTGGATACGCCGTTTTCCGACGC 859
        :   |||:::::::::      |||      ||||||::::::|||   :
    259 rLeuValIleHisHisAlaIleGluAspLeuIleArgLysTyrProAspH 276

    860 AGTATCTGTTTATGTACAACCGCTACAAA 888
        ::|||   :::   |||:::|||:::|||
    276 isTyrHisTrpSerTyrLysArgPheLys 285

seq_name: sp_bacteriap:Q9PDN1

seq_documentation_block:
ID   Q9PDN1      PRELIMINARY;      PRT;   316 AA.
AC   Q9PDN1;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   LIPID A BIOSYNTHESIS LAUROYL ACYLTRANSFERASE.
GN   XF1348.
OS   Xylella fastidiosa.
OC   Bacteria; Proteobacteria; gamma subdivision; Xanthomonas group;
OC   Xylella.
OX   NCBI_TaxID=2371;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=9A5C;
RX   MEDLINE=20365717; PubMed=10910347;
RA   Simpson A.J.G., Reinach F.C., Arruda P., Abreu F.A., Acencio M.,
RA   Alvarenga R., Alves L.M.C., Araya J.E., Baia G.S., Baptista C.S.,
RA   Barros M.H., Bonaccorsi E.D., Bordin S., Bove J.M., Briones M.R.S.,
RA   Bueno M.R.P., Camargo A.A., Camargo L.E.A., Carraro D.M., Carrer H.,
RA   Colauto N.B., Colombo C., Costa F.F., Costa M.C.R., Costa-Neto C.M.,
RA   Coutinho L.L., Cristofani M., Dias-Neto E., Docena C., El-Dorry H.,
RA   Facincani A.P., Ferreira A.J.S., Ferreira V.C.A., Ferro J.A.,
RA   Fraga J.S., Franca S.C., Franco M.C., Frohme M., Furlan L.R.,
RA   Garnier M., Goldman G.H., Goldman M.H.S., Gomes S.L., Gruber A.,
RA   Ho P.L., Hoheisel J.D., Junqueira M.L., Kemper E.L., Kitajima J.P.,
RA   Krieger J.E., Kuramae E.E., Laigret F., Lambais M.R., Leite L.C.C.,
RA   Lemos E.G.M., Lemos M.V.F., Lopes S.A., Lopes C.R., Machado J.A.,
RA   Machado M.A., Madeira A.M.B.N., Madeira H.M.F., Marino C.L.,
RA   Marques M.V., Martins E.A.L., Martins E.M.F., Matsukuma A.Y.,
RA   Menck C.F.M., Miracca E.C., Miyaki C.Y., Monteiro-Vitorello C.B.,
RA   Moon D.H., Nagai M.A., Nascimento A.L.T.O., Netto L.E.S.,
RA   Nhani A. Jr., Nobrega F.G., Nunes L.R., Oliveira M.A.,
RA   de Oliveira M.C., de Oliveira R.C., Palmieri D.A., Paris A.,
RA   Peixoto B.R., Pereira G.A.G., Pereira H.A. Jr., Pesquero J.B.,
RA   Quaggio R.B., Roberto P.G., Rodrigues V., de Rosa A.J.M.,
RA   de Rosa V.E. Jr., de Sa R.G., Santelli R.V., Sawasaki H.E.,
RA   da Silva A.C.R., da Silva A.M., da Silva F.R., Silva W.A. Jr.,
RA   da Silveira J.F., Silvestri M.L.Z., Siqueira W.J., de Souza A.A.,
RA   de Souza A.P., Terenzi M.F., Truffi D., Tsai S.M., Tsuhako M.H.,
RA   Vallada H., Van Sluys M.A., Verjovski-Almeida S., Vettore A.L.,
RA   Zago M.A., Zatz M., Meidanis J., Setubal J.C.;
RT   "The genome sequence of the plant pathogen Xylella fastidiosa.";
RL   Nature 406:151-159(2000).
DR   EMBL; AE003967; AAF84157.1; -.
KW   Complete proteome.
SQ   SEQUENCE   316 AA;  36286 MW;  8450EED03A405D3A CRC64;


alignment_scores:
            Quality:  263.50            Length:    286
              Ratio:  1.532               Gaps:      9
 Percent Similarity:  60.140   Percent Identity: 27.273

alignment_block:
 US-09-303-518D-569 x Q9PDN1   ..

 Align seg 1/1  to: Q9PDN1  from: 1  to: 316
```

```
    103 CTGCACACGCTGGGAAACCGGCTCGGACATCTGGCGTTTTACCTTTTAAA 152
        |||:::   :::::::::   :::|||||||:::   :::   :::|||||
     10 LeuTyrArgValAlaSerLeuValGlyHisIleProTrpProPheLeuLy 26

    153 GGAAGACCGCGCGCGCATCGTCGCC.................AATATGC 184
        |:::      :::   ::::::|||                 |||   |
     26 sArg...LeuSerAspLeuIleAlaTyrCysTrpLeuCysLeuAsnAlaA 42

    185 GTCAGGCA.................................GGCATG 198
        ||:::::                                     :::
     42 rgGluSerArgValThrGlnArgAsnLeuGluLeuThrTyrProGluLeu 58

    199 AATCCCGACCCCAAAACG......GTCAAAGCCGTTTTTGCGGAAACGGC 242
        :::|||:::      |||               ::::::   :::||||::
     59 SerProGlnGluHisThrArgLeuHisHisGlnIleIleTyrSerThrVa 75

    243 AAAAGGCGGTTTGGAACTTGCCCCCGCGTTTTTCAGAAAACCGGAAGACA 292
        ::::   ::::::|||:::          :::   :::          :::
     75 lArgGlnAlaPheGluMetLeuHisIleTrpThrHisProGluThrGluA 92

    293 TAGAAACAATGTTCAAAGCGGTACACGGCTGGGAACATGTGCAGCAGGCT 342
              :::   ::::::      ||||||   :::      :::   |||
     92 snLeuAlaArgLeuArgGluTyrHisGlyValAspLeuTyrSerAlaAla 108

    343 TTGGACAAACACGAAGGGCTGCTATTCATCACGCCGCACATCGGCAGCTA 392
        |||   :::::::::|||:::::::::::::::::::||||||:::|||:::::
    109 LeuAlaArgAsnArgGlyMetIleIleValValProHisPheGlyAsnTr 125

    393 CGATTTGGGCGGACGCTACATCAGCCAGCAGCTTCCGTTCCCGCTGACCG 442
        ::::|||      :::|||:::|||:::              |||||||||
    125 pGluLeuLeuLysGlnTyrLeuSerAspLeuAla......ProLeuThrL 140

    443 CCATGTACAAACCGCCGAAAATCAAAGCGATAGACAAAATCATGCAGGCG 492
          ||||||:::|||   :::      :::::::|||   :::::::|||
    140 euMetTyrArgProAlaAsnSerAlaValLeuAspGlyPheLeuGln... 155

    493 GGCAGGGTTCGCGGCAAAGGAAAAACCGCGCCTACCAGCATACAAGGG.. 540
           ||||||||||||   :::::::::      :::      :::|||
    156 ...ArgValArgGlyGlyAsnAsnValHisGlnValArgAlaGluGlyMe 171

    541 ....GTCAAACAAATCATCAAAGCCCTGCGTTCGGGCGAAGCAACCATCG 586
            |||:::::::::::|||:::|||:::::::|||   :::      :
    171 tAlaValArgHisLeuPheLysValLeuLysAsnGlyGlyThrIleGlyI 188

    587 TCCTGCCCGACCACGTCCCCTCCCCTCAAGAAGGCGGGGAAGGCGTATGG 636
        ::|||||||||:::   |||      :::   |||||||||||:::
    188 leLeuProAspGlnGlnPro.........LysIleGlyGluGlyValPhe 201

    637 GTGGATTTCTTCGGCAAACCTGCCTATACCATGACGCTGGCGGCAAAATT 686
        :::   |||||||||      |||   |||||||||||:::   :::||
    202 AlaProPhePheGlyIleGluAlaLeuThrMetThrLeuValGlnArgLe 218

    687 GGCACACGTCAAAGGCGTGAAAACCCTGTTTTTCTGCTGCGAACGCCTGC 736
        ||||:::      ::::::   :::|||:::      ||||||||||:::
    218 uAlaGluArgThrSerAlaThrValLeuTyrAlaTrpCysGluArgIleS 235

    737 CTGGCGGACAAGGTTTCGATTTGCACATCCGCCCCGTCCAAGGGGAATTG 786
                    |||   ||||||::::::   ::::::      :::
    235 erProHisLeuGluPheAlaLeuHisMetGluHisAlaAspProAlaVal 251

    787 ...AACGGCGACAAAGCCCATGATGCCGCCGTGTTCAACCGCAATGCCGA 833
           :::   |||           |||::::::::::|||   :::   ||
    252 AlaAspProAspProLeuIleAlaAlaThrAlaLeuAsnAlaGlyIleGl 268

    834 ATATTGGATACGCCGTTTTCCGACGCAGTATCTGTTTATGTACAACCGCT 883
        |         ||||||   ||||||||||||   :::   |||:::||||
    268 uArgIleAlaArgArgAspProThrGlnTyrGlnTrpThrTyrLysArgT 285

    884 ACAAAATG 891
        ||:::::::
    285 yrSerLeu 287

seq_name: sp_bacteriap:Q9I7B5
```

```
seq_documentation_block:
ID   Q9I7B5      PRELIMINARY;      PRT;   295 AA.
AC   Q9I7B5;
DT   01-MAR-2001 (TrEMBLrel. 16, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-OCT-2001 (TrEMBLrel. 18, Last annotation update)
DE   PROBABLE 2-OH-LAUROYLTRANSFERASE.
GN   PA0011.
OS   Pseudomonas aeruginosa.
OC   Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
OC   Pseudomonas.
OX   NCBI_TaxID=287;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=ATCC 15692 / PAO1;
RX   MEDLINE=20437337; PubMed=10984043;
RA   Stover C.K., Pham X.-Q.T., Erwin A.L., Mizoguchi S.D., Warrener P.,
RA   Hickey M.J., Brinkman F.S.L., Hufnagle W.O., Kowalik D.J., Lagrou M.,
RA   Garber R.L., Goltry L., Tolentino E., Westbrock-Wadman S., Yuan Y.,
RA   Brody L.L., Coulter S.N., Folger K.R., Kas A., Larbig K., Lim R.M.,
RA   Smith K.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.,
RT   "Complete genome sequence of Pseudomonas aeruginosa PAO1, an
RT   opportunistic pathogen.";
RL   Nature 406:959-964(2000).
DR   EMBL; AE004441; AAG03401.1; -.
KW   Transferase; Complete proteome.
SQ   SEQUENCE   295 AA;  33001 MW;  D6D1710EC6AD42FF CRC64;


alignment_scores:
            Quality:  259.00              Length:    286
              Ratio:  1.463                Gaps:      7
 Percent Similarity:  61.888   Percent Identity: 25.874

alignment_block:
 US-09-303-518D-569 x Q9I7B5   ..

 Align seg 1/1  to: Q9I7B5  from: 1  to: 295

      55 CTGTTGACCGCCCTGCTCAAATGCCTCTCCCTGCTGCCGCTTTCCTGTCT 104
         |||:::::::::   |||:::   ::::::|||||||||      :::::
       8 LeuValValGlyAlaLeuArgLeuPheAlaLeuLeuProTrpArgAlaVa 24

     105 GCACACGCTGGGAAACCGGCTCGGACATCTGGCGTTTTACCTTTTAAAGG 154
         ::::   :::|||      :::|||   |||   :::   |||
      24 lGlnGlyValGlyAlaGlyIleGlyTrpLeuMetTrpLysLeu......P 39

     155 AAGACCGCGCGCGCATCGTCGCCAATATGCGTCAGGCAGGCATGAATCCC 204
           :::|||:::|||   |||:::::::::::::   :::          |||
      39 roAsnArgSerArgGluValValArgIleAsnLeuSerLysCysPhePro 55

     205 GAC.................CCCAAAACGGTCAAAGCCGTTTTTGCGGA 236
         :::                |||   |||         :::   ::
      56 GluLeuSerGluThrGluLeuGluLysLeuValGlyGlnSerLeuMetAs 72

     237 AACGGCAAAAGGCGGTTTGGAACTTGCCCCCGCGTTTTTCAGAAAACCGG 286
         :   ::::::          |||   |||   |||::::::      ||||
      72 pIleGlyArgThrLeuThrGluSerAlaCysAlaTrpIleTrpProProG 89

     287 AAGACATAGAAACAATGTTCAAAGCGGTACACGGCTGGGAACATGTGCAG 336
         ||               ::::::   |||:::|||   |||   ::::::
      89 luLysSerLeuArgTyrIleArgGluValGluGlyMetGluValLeuGlu 105

     337 CAGGCTTTGGACAAACACGAAGGGCTGCTATTCATCACGCCGCACATCGG 386
         :::||||||   :::   :::||||||:::   ||||||   |||:::||
     106 GluAlaLeuAlaSerGlyAspGlyLeuValGlyIleThrSerHisLeuGl 122

     387 CAGCTACGAT...TTGGGCGGACGCTACATCAGCCAGCAGCTTCCGTTCC 433
         |:::::::::   |||:::      |||   |||         |||:::
     122 yAsnTrpGluValLeuAsnHisPheTyrCysSerTyrAlaLysProIle. 138

     434 CGCTGACCGCCATGTACAAACCGCCGAAAATCAAAGCGATAGACAAAATC 483
                 :::|||:::|||||||||:::||||||:::|||::::::
     139 ........IlePheTyrArgProProLysLeuLysAlaValAspGluLeu 152

     484 ATGCAGGCGGGCAGGGTTCGCGGCAAAGGAAAAACCGCGCCTACCAGCAT 533
         ::::::      ||||||:::      :::::::::||||||::::::
     153 LeuLysLysGlnArgValGlnLeuGlyAsnArgValAlaProSerThrPr 169

     534 ACAAGGGGTCAAACAAATCATCAAAGCCCTGCGTTCGGGCGAAGCAACCA 583
         :::|||:::   ::::::||||||   ::::::::|||   ::::::
     169 oGluGlyIleLeuSerValIleLysGluValLysLysGlyGlyCysValG 186

     584 TCGTCCTGCCCGACCACGTCCCCTCCCCTCAAGAAGGCGGGGAAGGCGTA 633
           :::      |||      |||:::|||   :::   :::   |||:::
     186 lyIleProAlaAsp......ProGluProAlaArgThrAla...GlyLeu 199

     634 TGGGTGGATTTCTTCGGCAAACCTGCCTATACCATGACGCTGGCGGCAAA 683
         :::|||   ::::::|||      |||                ::::::   ::
     200 PheValProTyrLeuGlyThrThrAlaLeuIleSerLysPheValProGl 216

     684 ATTGGCACACGTCAAAGGCGTGAAAACCCTGTTTTTCTGCTGCGAACGCC 733
         :|||                ::::::   :::|||||   :::   ||||
     216 nLeuLeuSerArgGlyLysAlaArgGlyValPhePheHisAlaValArgL 233

     734 TGCCTGGCGGACAAGGTTTCGATTTGCACATCCGCCCCGTCCAAGGGGAA 783
         |||||   |||:::|||:::   :::   ::::::   :::   ::::::
     233 euProAspGlySerGlyTyrLysValIleLeuGluAlaAlaProAlaAsp 249

     784 TTG...AACGGCGACAAAGCCCATGATGCCGCCGTGTTCAACCGCAATGC 830
         :::   :::   |||         ::::::|||:::::::::|||:::
     250 MetTyrAspLysAspLeuGluValSerValAlaAlaMetSerArgGluLe 266

     831 CGAATATTGGATACGCCGTTTTCCGACGCAGTATCTGTTTATGTACAACC 880
               ::::::|||   :::|||:::|||||||:::::      :::|
     266 uAlaLysTyrValArgAlaTyrProSerGlnTyrMetTrpSerMetLysA 283

     881 GCTACAAA 888
         ||:::|||
     283 rgPheLys 285

seq_name: sp_bacteriap:Q9HYZ8

seq_documentation_block:
ID   Q9HYZ8      PRELIMINARY;      PRT;   312 AA.
AC   Q9HYZ8;
DT   01-MAR-2001 (TrEMBLrel. 16, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-OCT-2001 (TrEMBLrel. 18, Last annotation update)
DE   PROBABLE LAUROYL ACYLTRANSFERASE.
GN   PA3242.
OS   Pseudomonas aeruginosa.
OC   Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
OC   Pseudomonas.
OX   NCBI_TaxID=287;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=ATCC 15692 / PAO1;
RX   MEDLINE=20437337; PubMed=10984043;
RA   Stover C.K., Pham X.-Q.T., Erwin A.L., Mizoguchi S.D., Warrener P.,
RA   Hickey M.J., Brinkman F.S.L., Hufnagle W.O., Kowalik D.J., Lagrou M.,
RA   Garber R.L., Goltry L., Tolentino E., Westbrock-Wadman S., Yuan Y.,
RA   Brody L.L., Coulter S.N., Folger K.R., Kas A., Larbig K., Lim R.M.,
RA   Smith K.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.,
RA   Reizer J., Saier M.H., Hancock R.E.W., Lory S., Olson M.V.;
RT   "Complete genome sequence of Pseudomonas aeruginosa PAO1, an
RT   opportunistic pathogen.";
RL   Nature 406:959-964(2000).
DR   EMBL; AE004747; AAG06630.1; -.
KW   Transferase; Acyltransferase; Complete proteome.
SQ   SEQUENCE   312 AA;  36063 MW;  0EDE771D75F3C186 CRC64;


alignment_scores:
```

```
          Quality:  228.50              Length:    281
            Ratio:  1.336                 Gaps:      6
Percent Similarity:  60.854   Percent Identity: 26.335

alignment_block:
 US-09-303-518D-569 x Q9HYZ8   ..

 Align seg 1/1  to: Q9HYZ8  from: 1  to: 312

      67 CTGCTCAAATGCCTCTCCCTGCTGCCGCTTTCCTGTCTGCACACGCTGGG 116
         ||||||      :::      ||||||         |||      |||||
      24 LeuLeuTrpLeuValValGlnLeuProTyrProValLeuLeuMetLeuGl 40

     117 AAACCGGCTCGGACATCTGGCGTTTTACCTTTTAAAGGAAGACCGCGCGC 166
         |:::   ||||||   |||   :::   |||:::   :::   |||
      40 yArgGlyLeuGlyAlaLeuMetTyrArgLeuValGlySerArgArgGluI 57

     167 GCATCGTCGCCAATATGCGTCAGGCAGGCATGAATCCCGACCCC..... 210
             :::   |||::::::   :::      :::   :::|||
      57 leAlaAlaArgAsnLeuGluLeuCysPheProGluLysSerProAlaGlu 73

     211 ......AAAACGGTCAAAGCCGTTTTTGCGGAAACGGCAAAAGGCGGTTT 254
               :::   :::|||      ||||||::::::::::   :::   ::
      74 ArgGluArgLeuLeuLysGluAsnPheAlaSerSerGlyIleAlaPhePh 90

     255 GGAACTTGCCCCCGCGTTTTTCAGAAAACCGGAAGACATAGAAACAATGT 304
         :|||:::|||   :::::::::      |||:::          :::   :
      90 eGluMetAlaMetSerTrpTrp...TrpProLysAlaArgLeuAlaArgL 106

     305 TCAAAGCGGTACACGGCTGGGAACATGTGCAGCAGGCTTTGGACAAACAC 354
         ::      ::::::|||   ||||||::::::::::|||      :::
     106 euAlaHisIleGluGlyLeuGluHisLeuArgGluAlaGlnAlaGlnGly 122

     355 GAAGGGCTGCTATTCATCACGCCGCACATCGGCAGCTACGATTTGGGCGG 404
         ||||||:::::::::::::::   |||:::   :::   ::::::|||::
     123 GluGlyValIleLeuMetAlaLeuHisPheThrThrLeuGluIleGlyAl 139

     405 ACGCTACATCAGCCAGCAGCTTCCGTTCCCGCTGACCGCCATGTACAAAC 454
         :      ::::::|||            :::   :::|||||||:::
     139 aAlaLeuLeuGlyGln......ValHisThrIleAspGlyMetTyrArgG 154

     455 CGCCGAAAATCAAAGCGATAGACAAAATCATGCAGGCGGGCAGGGTTCGC 504
                    :::::::|||   :::::::::   ||||||   |||
     154 luHisAspAsnProValPheAspTyrValGlnArgArgGlyArgGluArg 170

     505 GGCAAAGGAAAAACCGCGCCTACCAGCATACAAGGGGTCAAACAAATCAT 554
         :::      ::::::      :::   :::   |||::::::::::::
     171 HisAsnLeuAspAlaThrAlaIleGluArgGluAspValArgSerMetLe 187

     555 CAAAGCCCTGCGTTCGGGCGAAGCAACCATCGTCCTGCCCGACCACGTCC 604
         :|||:::||||||:::|||:::|||            ||||||
     187 uLysValLeuArgGlyGlyArgAlaIleTrpTyrAlaProAsp..... 201

     605 CCTCCCCTCAAGAAGGCGGG......GAAGGCGTATGGGTGGATTTCTTC 648
               |||:::   |||      ::::::::::::|||   :::||||
     202 ........GlnAspTyrGlyAlaLysGlnSerLeuPheValProLeuPhe 215

     649 GGCAAACCTGCCTATACCATGACGCTGGCGGCAAAATTGGCACACGTCAA 698
         |||   ||||||   |||:::|||   ::::::|||:::|||::::::
     216 GlyIleProAlaAlaThrValThrAlaThrThrLysPheAlaArgLeuGl 232

     699 AGGCGTGAAAACCCTGTTTTTCTGCTGCGAACGCCTGCCTGGCGGACAAG 748
             ::::::::::|||   |||      :::||||||      |||:::|
     232 yArgAlaArgValLeuProPheThrGlnSerArgLeuAlaAspGlySerG 249

     749 GTTTCGATTTGCACATCCGCCCCGTCCAAGGGGAATTGAACGGCGACAAA 798
         ||:::   |||   |||:::|||         ::::::   |||::::::
     249 lyTyrArgLeuThrIleHisProProLeuGluAspPheProGlyGluSer 265

     799 GCCCAT...GATGCCGCCGTGTTCAACCGCAATGCCGAATATTGGATACG 845
             :::   |||:::      :::|||:::   :::|||      :::||
     266 GluGluAlaAspCysLeuArgIleAsnGlnTrpValGluArgCysValAr 282

     846 CCGTTTTCCGACGCAGTATCTGTTTATGTACAACCGCTACAAA 888
         |:::   |||   |||||||||:::   ::::::|||:::|||
     282 gGlnGlnProGluGlnTyrLeuTrpAlaHisArgArgPheLys 296

seq_name: sp_bacteria:O33437

seq_documentation_block:
ID   O33437      PRELIMINARY;      PRT;   253 AA.
AC   O33437;
DT   01-JAN-1998 (TrEMBLrel. 05, Created)
DT   01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   HTRB PROTEIN (FRAGMENT).
GN   HTRB.
OS   Pseudomonas fluorescens.
OC   Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
OC   Pseudomonas.
OX   NCBI_TaxID=294;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=WCS365;
RX   MEDLINE=98340543; PubMed=9675892;
RA   Dekkers L.C., van der Bij A.J., Mulders H.M., Phoelich C.C.,
RA   Wentwoord A.R., Glandorf D.C.M., Wijffelman C.A., Lugtenberg B.J.J.;
RT   "Role of the O-antigen of lipopolysaccharide, and possible roles of
RT   growth rate and NADH:ubiquinone oxidoreductase (nuo) in competitive
RT   tomato root-tip colonization by Pseudomonas fluorescens WCS365.";
RL   Mol. Plant Microbe Interact. 11:763-771(1998).
DR   EMBL; Y14568; CAA74897.1; -.
FT   NON_TER     253    253
SQ   SEQUENCE   253 AA;  27634 MW;  F12E003B10C8223D CRC64;


alignment_scores:
          Quality:  221.50              Length:    239
            Ratio:  1.467                 Gaps:      7
Percent Similarity:  63.180   Percent Identity: 28.870

alignment_block:
 US-09-303-518D-569 x O33437   ..

 Align seg 1/1  to: O33437  from: 1  to: 253

      55 CTGTTGACCGCCCTGCTCAAATGCCTCTCCCTGCTGCCGCTTTCCTGTCT 104
         ||||||::::::   |||:::   ::::::|||||||||      :::::
       8 LeuLeuValGlyAlaLeuArgLeuPheAlaLeuLeuProTrpArgAlaVa 24

     105 GCACACGCTGGGAAACCGGCTCGGACATCTGGCGTTTTACCTTTTAAAGG 154
         ::::::::::|||:::   :::|||   :::   :::   |||   ::::
      24 lGlnAlaValGlySerAlaIleGlyTrpPheMetTrpLysLeuProAsnA 41

     155 AAGACCGCGCGCGCATCGTCGCCAATATGCGTCAGGCAGGCATGAATCCC 204
         :::::|||      :::      |||:::   ::::::      :::
      41 rgSerArgAspValValArgIleAsnLeuAlaLysCysPheProGluMet 57

     205 GACCCC............AAAACGGTCAAAGCCGTTTTTGCGGAAACGGC 242
         ||||||            :::   |||         :::   ::   ::
      58 AspProAlaGluArgGluArgLeuValGlyGlnSerLeuLysAspIleGl 74

     243 AAAAGGCGGTTTGGAACTTGCCCCCGCGTTTTTCAGAAAACCGGAAGACA 292
         :|||:::      |||   |||   |||::::::         :::
      74 yLysSerLeuThrGluSerAlaCysAlaTrpIleTrpProAlaGlnArgS 91

     293 TAGAAACAATGTTCAAAGCGGTACACGGCTGGGAACATGTGCAGCAGGCT 342
                 :::   :::   |||:::|||   |||   ::::::::::|||
      91 erIleAspLeuValArgGluValGluGlyLeuGluValLeuLysGluAla 107

     343 TTGGACAAACACGAAGGGCTGCTATTCATCACGCCGCACATCGGCAGCTA 392
         |||   :::   :::|||:::::::   ||||||   |||:::|||:::::
     108 LeuAlaSerGlyLysGlyValValGlyIleThrSerHisLeuGlyAsnTr 124

     393 CGAT...TTGGGCGGACGCTACATCAGCCAGCAGCTTCCGTTCCCGCTGA 439
         ::::   |||:::      |||   ||||||      |||:::
```

```
 124 pGluValLeuAsnHisPheTyrCysSerGlnCysLysProIle..... 138

 440 CCGCCATGTACAAACCGCCGAAAATCAAAGCGATAGACAAAATCATGCAG 489
     :::|||:::|||||||||:::||||||:::|||||||||||
 139 ..IlePheTyrArgProProLysLeuLysAlaValAspGluLeuLeuArg 154

 490 GCGGGCAGGGTTCGCGGCAAAGGAAAAACCGCGCCTACCAGCATACAAGG 539
         |||||||:::      :::|||:::|||   ::::::   :::||
 155 LysGlnArgValGlnLeuGlyAsnLysValAlaAlaSerThrLysGluGl 171

 540 GGTCAAACAAATCATCAAAGCCCTGCGTTCGGGCGAAGCAACCATCGTCC 589
     |:::   ::::::|||||||   :::|||:::|||      :::   :::
 171 yIleLeuSerValIleLysGluValArgLysGlyGlyGlnValGlyIleP 188

 590 TGCCCGACCACGTCCCCTCCCCTCAAGAAGGCGGGGAAGGCGTATGGGTG 639
         |||      |||:::|||   |||:::::   |||:::::|||
 188 roAlaAsp......ProGluProAlaGluSerAla...GlyIlePheVal 201

 640 GATTTCTTCGGCAAACCTGCCTATACCATGACGCTGGCGGCAAAA...TT 686
     ||||||:::      |||   |||      :::::   :::   ||
 202 ProPhePheAlaThrGlnAlaLeuThrSerLysPheValProAsnMetLe 218

 687 GGCACACGTCAAAGGCGTGAAAACCCTGTTTTTCTGCTGCGAACGCCTGC 736
     ||||      |||:::|||      :::|||:::   :::   |||||||
 218 uAlaGlyGlyLysAlaVal...GlyValPheLeuHisAlaLeuArgLeuP 234

 737 CTGGCGGACAAGGTTTC 753
     ||   |||:::|||:::
 234 roAspGlySerGlyTyr 239
```

seq_name: sp_bacteriap:Q9JYV2

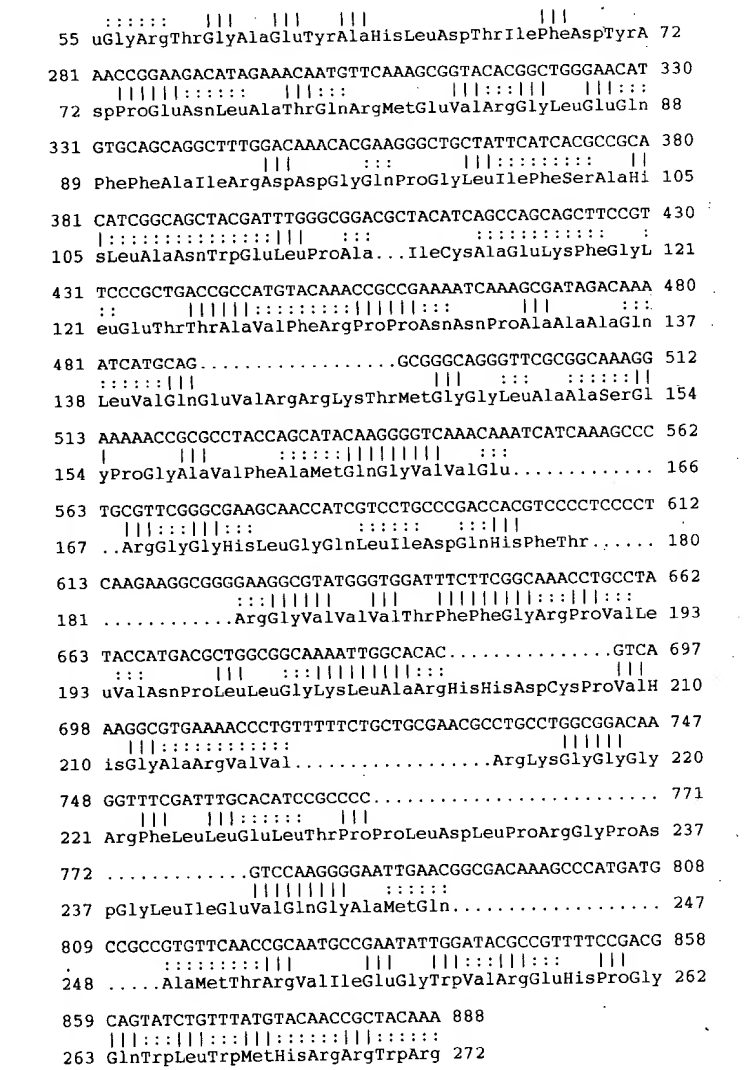
```
seq_documentation_block:
ID   Q9JYV2          PRELIMINARY;      PRT;   289 AA.
AC   Q9JYV2;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   HTRB/MSBB FAMILY PROTEIN.
GN   NMB1418.
OS   Neisseria meningitidis (serogroup B).
OC   Bacteria; Proteobacteria; beta subdivision; Neisseriaceae; Neisseria.
OX   NCBI_TaxID=491;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=MC58 / SEROGROUP B;
RX   MEDLINE=20175755; PubMed=10710307;
RA   Tettelin H., Saunders N.J., Heidelberg J., Jeffries A.C., Nelson K.E.,
RA   Eisen J.A., Ketchum K.A., Hood D.W., Peden J.F., Dodson R.J.,
RA   Nelson W.C., Gwinn M.L., DeBoy R., Peterson J.D., Hickey E.K.,
RA   Haft D.H., Salzberg S.L., White O., Fleischmann R.D., Dougherty B.A.,
RA   Mason T., Ciecko A., Parksey D.S., Blair E., Cittone H., Clark E.B.,
RA   Cotton M.D., Utterback T.R., Khouri H., Qin H., Vamathevan J.,
RA   Gill J., Scarlato V., Masignani V., Pizza M., Grandi G., Sun L.,
RA   Smith H.O., Fraser C.M., Moxon E.R., Rappuoli R., Venter J.C.;
RT   "Complete genome sequence of Neisseria meningitidis serogroup B strain
RT   MC58.";
RL   Science 287:1809-1815(2000).
DR   EMBL; AE002491; AAF41779.1; -.
DR   TIGR; NMB1418; -.
KW   Complete proteome.
SQ   SEQUENCE   289 AA;  33843 MW;  3BF100F050576512 CRC64;
```

```
alignment_scores:
            Quality:  204.50
              Ratio:  1.225              Length:    286
Percent Similarity:  58.392   Percent Identity: 26.923
                                          Gaps:     11

alignment_block:
US-09-303-518D-569 x Q9JYV2

Align seg 1/1  to: Q9JYV2  from: 1  to: 289
```

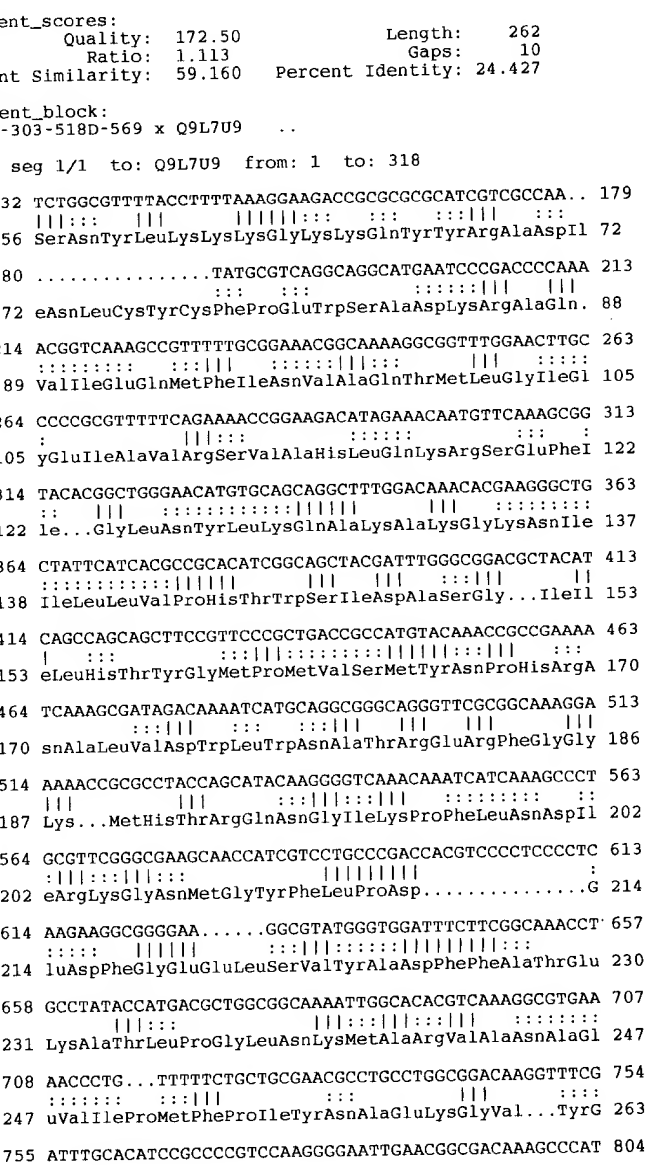
```
  67 CTGCTCAAATGCCTCTCCCTGCTGCCGCTTTCCTGTCTGCACACGCTGGG 116
     :::|||      |||:::::|||||||:::::::  ||||||  :::::
   7 ValLeuTyrValLeuGlnPheLeuProPheAlaLeuLeuHisLysIleAl 23

 117 AAACCGGCTCGGACATCTGGCGTTTTACCTTTTAAAGGAAGACCGCGCGC 166
     ::::      |||  ||||||  |||||||||  :::  |||  |
  23 aAspLeuThrGlyLeuLeuAla...TyrLeuLeuValLysProArgArgA 39

 167 GCATC............GTCGCCAATATGCGTCAGGCAGGCATGAATCCC 204
     |||||            :::|||:::      :::        :::
  39 rgIleGlyGluIleAsnLeuAlaLysCysPheSerGluTrpSerGluGlu 55

 205 GACCCCAAAACGGTC...AAAGCCGTTTTTGCGGAAACGGCAAAAGGCGG 251
         |||||||||  |||      |||   :::  ||||||
  56 LysArgLysThrValLeuLysGlnHisPheLysHisMetAlaLysLeuMe 72

 252 TTTGGAACTTGCCCCCGCGTTTTTCAGAAAACCGGAAGACATAGAAACAA 301
     ||||||  :::      :::::                ::::::::::
  72 tLeuGluTyrGlyLeuTyrTrpTyrAlaProAlaGlyArgLeuLysSerL 89

 302 TGTTCAAAGCGGTACACGGCTGGGAACATGTGCAGCAGGCTTTGGACAAA 351
     ::  :::      :::::   ::::::::::::::::||||||
  89 euValArg...TyrArgAsnLysHisTyrLeuAspAspAlaLeuAlaAla 104

 352 CACGAAGGGCTGCTATTCATCACGCCGCACATCGGCAGCTACGATTTGGG 401
         |||   ::::::::::::   |||||||::   ::::::::::::
 105 GlyGluLysValIleIleLeuTyrProHisPheThrAlaPheGluMetAl 121

 402 CGGACGCTACATCAGCCAGCAGCTTCCGTTCCCGCTGACCGCCATGTACA 451
     :      ::::::|||::::::      ||||||   :::|||||||:
 121 aValTyrAlaLeuAsnGlnAspIle......ProLeuIleSerMetTyrS 136

 452 AACCGCCGAAAATCAAAGCGATAGACAAAATCATGCAGGCGGGCAGGGTT 501
     ::      |||  |||   :::|||:::      :::      ||||||
 136 erHisGlnLysAsnLysIleLeuAspGluGlnIleLeuLysGlyArgAsn 152

 502 CGCGGCAAAGGAAAAACCGCGCCTACCAGCATACAAGGGGTCAAACAAAT 551
     |||      :::                          :::|||:::::   ::
 153 ArgTyrHisAsnValPheLeuIleGlyArgThrGluGlyLeuArgAlaLe 169

 552 CATCAAAGCCCTGCGTTCGGGCGAAGCAACC...ATCGTCCTGCCCGACC 598
     :::::|||   :::|||:::::::::|||      :::   |||||||||
 169 uValLysGlnPheArgLysSerSerAlaProPheLeuTyrLeuProAsp. 185

 599 ACGTCCCCTCCCCTCAAGAAGGCGGG......GAAGGCGTATGGGTGGAT 642
                 |||:::   |||      ::::::|||:::||||||
 186 .............GlnAspPheGlyArgAsnAspSerValPheValAsp 197

 643 TTCTTCGGCAAACCTGCCTATACCATGACGCTGGCGGCAAAATTGGCACA 692
     |||||||||      :::   |||:::|||      ::::::::|||
 198 PhePheGlyIleGlnThrAlaThrIleThrGlyLeuSerArgIleAlaAl 214

 693 CGTCAAAGGCGTGAAAACCCTGTTTTTCTGCTGCGAACGCCTGCCT.... 738
     :::   ::::::|||:::::::                  :::|||
 214 aLeuAlaAsnAlaLysValIle.............ProAlaIleProValA 227

 739 .....GGCGGACAAGGTTTCGATTTGCACATCCGCCCCGTCCAAGGGGAA 783
         :::   :::          |||||||:::   |||:::      :::
 227 rgGluAlaAspAsnThrValThrLeuHisPheTyrProAlaTrpLysSer 243

 784 TTGAACGGCGACAAAGCCCAT...GATGCCGCCGTGTTCAACCGCAATGC 830
     :::   |||:::   |||      ||||||      :::||||||
 244 PheProGlyGluAspAlaLysAlaAspAlaGlnArgMetAsnArgPheIl 260

 831 CGAATATTGGATACGCCGTTTTCCGACGCAGTATCTGTTTATGTACAACC 880
     |||      :::|||:::   |||   ||||||:::::::::::::::|
 260 eGluAspArgValArgGluHisProGluGlnTyrPheTrpLeuHisLysA 277

 881 GCTACAAA 888
     ||:::|||
 277 rgPheLys 279
```

```
seq_name: sp_bacteriap:Q9JTU4

seq_documentation_block:
ID   Q9JTU4      PRELIMINARY;      PRT;   289 AA.
AC   Q9JTU4;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   PUTATIVE ACETYLTRANSFERASE.
GN   NMA1630.
OS   Neisseria meningitidis (serogroup A).
OC   Bacteria; Proteobacteria; beta subdivision; Neisseriaceae; Neisseria.
OX   NCBI_TaxID=65699;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Z2491 / SEROGROUP A / SEROTYPE 4A;
RX   MEDLINE=20222556; PubMed=10761919;
RA   Parkhill J., Achtman M., James K.D., Bentley S.D., Churcher C.,
RA   Klee S.R., Morelli G., Basham D., Brown D., Chillingworth T.,
RA   Davies R.M., Davis P., Devlin K., Feltwell T., Hamlin N., Holroyd S.,
RA   Jagels K., Leather S., Moule S., Mungall K., Quail M.A.,
RA   Rajandream M.A., Rutherford K.M., Simmonds M., Skelton J.,
RA   Whitehead S., Spratt B.G., Barrell B.G.;
RT   "Complete DNA sequence of a serogroup A strain of Neisseria
RT   meningitidis Z2491.";
RL   Nature 404:502-506(2000).
DR   EMBL; AL162756; CAB84858.1; -.
KW   Transferase; Complete proteome.
SQ   SEQUENCE   289 AA;  33867 MW;  4FAE453B5A632C1D CRC64;


alignment_scores:
            Quality:  203.50              Length:    285
              Ratio:  1.219                Gaps:      9
 Percent Similarity:  58.596   Percent Identity: 25.965

alignment_block:
 US-09-303-518D-569 x Q9JTU4   ..

 Align seg 1/1  to: Q9JTU4  from: 1  to: 289

      67 CTGCTCAAATGCCTCTCCCTGCTGCCGCTTTCCTGTCTGCACACGCTGGG 116
         :::|||      |||::::::||||||:::::   ||||||   |||::
       7 ValLeuTyrValLeuGlnPheLeuProPheAlaLeuLeuHisLysLeuAl 23

     117 AAACCGGCTCGGACATCTGGCGTTTTACCTTTTAAAGGAAGACCGCGCGC 166
         ::::      |||   ||||||:::   |||:::|||      |||
      23 aAspLeuThrGlyLeuLeuAlaTyrLeuLeuValLysProArgArgArgI 40

     167 GCATCGTCGCCAATATGCGTCAGGCAGGCATGAATCCCGACCCCAAAACG 216
                 |||:::  ::::::      :::   |||   |||
      40 leGlyGluIleAsnLeuAlaLysCysPheProGluTrpAspGlyLysLys 56

     217 GTCAAAGCCGTT...........TTTGCGGAAACGGCAAAAGGCGGTTT 254
            |||:::|||            |||   :::   ||||||      ||
      57 ArgLysThrValLeuLysGlnHisPheLysHisMetAlaLysLeuMetLe 73

     255 GGAACTTGCCCCCGCGTTTTTCAGAAAACCGGAAGACATAGAAACAATGT 304
         ||||   :::     ::::::                :::::::::::
      73 uGluTyrGlyLeuTyrTrpTyrAlaProAlaGlyArgLeuLysSerLeuV 90

     305 TCAAAGCGGTACACGGCTGGGAACATGTGCAGCAGGCTTTGGACAAACAC 354
            :::      ::::::   ::::::::::::::||||||
      90 alArg...TyrArgAsnLysHisTyrLeuAspAspAlaLeuAlaAlaGly 105

     355 GAAGGGCTGCTATTCATCACGCCGCACATCGGCAGCTACGATTTGGGCGG 404
         |||   ::::::::::   |||||||:::   ::::::::::::::
     106 GluLysValIleIleLeuTyrProHisPheThrAlaPheGluMetAlaVa 122

     405 ACGCTACATCAGCCAGCAGCTTCCGTTCCCGCTGACCGCCATGTACAAAC 454
            ::::::|||::::::      ||||||   :::|||||||:::
     122 lTyrAlaLeuAsnGlnAspVal......ProLeuIleSerMetTyrSerH 137

     455 CGCCGAAAATCAAAGCGATAGACAAAATCATGCAGGCGGGCAGGGTTCGC 504
            |||   |||   :::|||:::   :::         ||||||   |||
     137 isGlnLysAsnLysIleLeuAspGluGlnIleLeuLysGlyArgAsnArg 153

     505 GGCAAAGGAAAAACCGCGCCTACCAGCATACAAGGGGTCAAACAAATCAT 554
            :::                              :::|||::::::   :::::
     154 TyrHisAsnValPheLeuIleGlyArgThrGluGlyLeuArgAlaLeuVa 170

     555 CAAAGCCCTGCGTTCGGGCGAAGCAACC...ATCGTCCTGCCCGACCACG 601
         :|||   :::|||::::::::::|||      :::   |||||||||
     170 lLysGlnPheArgLysSerSerAlaProPheLeuTyrLeuProAsp.... 185

     602 TCCCCTCCCCTCAAGAAGGCGGG......GAAGGCGTATGGGTGGATTTC 645
                 |||:::   |||     ::::::|||:::|||||||||
     186 ..........GlnAspPheGlyArgAsnAspSerValPheValAspPhe 198

     646 TTCGGCAAACCTGCCTATACCATGACGCTGGCGGCAAAATTGGCACACGT 695
         ||||||      :::   |||:::|||      :::::::::|||   ::
     199 PheGlyIleArgThrAlaThrIleThrGlyLeuSerArgIleAlaAlaLe 215

     696 CAAAGGCGTGAAAACCCTGTTTTTCTGCTGCGAACGCCTGCCT..... 738
         :   ::::::|||:::::::                  :::|||
     215 uAlaAsnAlaLysValIle...........ProAlaIleProValArgG 228

     739 ..GGCGGACAAGGTTTCGATTTGCACATCCGCCCCGTCCAAGGGGAATTG 786
            :::   :::          |||||||:::   |||:::      ::::::
     228 luAlaAspAsnThrValThrLeuHisPheTyrProAlaTrpGluSerPhe 244

     787 AACGGCGACAAAGCCCAT...GATGCCGCCGTGTTCAACCGCAATGCCGA 833
            ::::::   |||:::   ||||||          :::|||||      ||
     245 ProSerGluAspAlaGlnAlaAspAlaGlnArgMetAsnArgPheIleGl 261

     834 ATATTGGATACGCCGTTTTCCGACGCAGTATCTGTTTATGTACAACCGCT 883
         |      :::|||:::   |||   ||||||:::::::::::::::|||:
     261 uGluArgValArgGluHisProGluGlnTyrPheTrpLeuHisLysArgP 278

     884 ACAAA 888
         ::|||
     278 heLys 279

seq_name: sp_bacteriap:Q9PH42

seq_documentation_block:
ID   Q9PH42      PRELIMINARY;      PRT;   279 AA.
AC   Q9PH42;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   LIPID A BIOSYNTHESIS LAUROYL ACYLTRANSFERASE.
GN   XF0104.
OS   Xylella fastidiosa.
OC   Bacteria; Proteobacteria; gamma subdivision; Xanthomonas group;
OC   Xylella.
OX   NCBI_TaxID=2371;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=9A5C;
RX   MEDLINE=20365717; PubMed=10910347;
RA   Simpson A.J.G., Reinach F.C., Arruda P., Abreu F.A., Acencio M.,
RA   Alvarenga R., Alves L.M.C., Araya J.E., Baia G.S., Baptista C.S.,
RA   Barros M.H., Bonaccorsi E.D., Bordin S., Bove J.M., Briones M.R.S.,
RA   Bueno M.R.P., Camargo A.A., Camargo L.E.A., Carraro D.M., Carrer H.,
RA   Colauto N.B., Colombo C., Costa F.F., Costa M.C.R., Costa-Neto C.M.,
RA   Coutinho L.L., Cristofani M., Dias-Neto E., Docena C., El-Dorry H.,
RA   Facincani A.P., Ferreira A.J.S., Ferreira V.C.A., Ferro J.A.,
RA   Fraga J.S., Franca S.C., Franco M.C., Frohme M., Furlan L.R.,
RA   Garnier M., Goldman G.H., Goldman M.H.S., Gomes S.L., Gruber A.,
RA   Ho P.L., Hoheisel J.D., Junqueira M.L., Kemper E.L., Kitajima J.P.,
RA   Krieger J.E., Kuramae E.E., Laigret F., Lambais M.R., Leite L.C.C.,
RA   Lemos E.G.M., Lemos M.V.F., Lopes S.A., Lopes C.R., Machado J.A.,
RA   Machado M.A., Madeira A.M.B.N., Madeira H.M.F., Marino C.L.,
RA   Marques M.V., Martins E.A.L., Martins E.M.F., Matsukuma A.Y.,
RA   Menck C.F.M., Miracca E.C., Miyaki C.Y., Monteiro-Vitorello C.B.,
```

```
RA   Moon D.H., Nagai M.A., Nascimento A.L.T.O., Netto L.E.S.,
RA   Nhani A. Jr., Nobrega F.G., Nunes L.R., Oliveira M.A.,
RA   de Oliveira M.C., de Oliveira R.C., Palmieri D.A., Paris A.,
RA   Peixoto B.R., Pereira G.A.G., Pereira H.A. Jr., Pesquero J.B.,
RA   Quaggio R.B., Roberto P.G., Rodrigues V., de Rosa A.J.M.,
RA   de Rosa V.E. Jr., de Sa R.G., Santelli R.V., Sawasaki H.E.,
RA   da Silva A.C.R., da Silva A.M., da Silva F.R., Silva W.A. Jr.,
RA   da Silveira J.F., Silvestri M.L.Z., Siqueira W.J., de Souza A.A.,
RA   de Souza A.P., Terenzi M.F., Truffi D., Tsai S.M., Tsuhako M.H.,
RA   Vallada H., Van Sluys M.A., Verjovski-Almeida S., Vettore A.L.,
RA   Zago M.A., Zatz M., Meidanis J., Setubal J.C.;
RT   "The genome sequence of the plant pathogen Xylella fastidiosa.";
RL   Nature 406:151-159(2000).
DR   EMBL; AE003864; AAF82917.1; -.
KW   Complete proteome.
SQ   SEQUENCE   279 AA;  32072 MW;  450EC409876A7791 CRC64;


alignment_scores:
           Quality:  196.50              Length:    275
             Ratio:  1.244                Gaps:     10
 Percent Similarity:  57.455   Percent Identity: 27.273

alignment_block:
 US-09-303-518D-569 x Q9PH42   ..

 Align seg 1/1  to: Q9PH42  from: 1  to: 279

      88 CTGCCGCTTTCCTGTCTGCACACGCTGGGAAACCGGCTCGGACATCTGGC 137
         ||||||   :::       :::   |||||||:::   :::|||   :::||
       7 LeuProTrpGlyLeuGlnArgArgLeuGlyArgValIleGlyAlaMetAl 23

     138 GTTTTACCTTTTAAAGGAAGACCGCGCGCGCATCGTCGCCAATATGCGTC 187
         |:::   |||:::         |||:::          :::|||:::|||
      23 aLeuArgLeuValGlyThrArgArgCysAlaAlaGluValAsnLeuArgL 40

     188 AGGCAGGCATGAATCCCGAC.............CCCAAAACGGTCAAAGCC 225
            :::      :::   |||                |||   :::
      40 euCysPheProThrHisAspGluAlaTrpArgLysLysLeuLeuHisGlu 56

     226 GTTTTTGCGGAAACGGCAAAAGGCGGTTTGGAACTTGCCCCCGCGTTTTT 275
         |||         :::   |||   :::|||:::|||   |||:::::
      57 AsnPheAspAlaLeuGlyValGlyLeuPheGluPheAlaArgAlaTrpTr 73

     276 CAGAAAACCGGAAGACATAGAAACAATGTTCAAAGCGGTACAC......G 319
         :            :::|||:::::::::      :::::::|||      |
      73 p.........GlySerIleAspValIleArgAlaGlyIleHisValGluG 87

     320 GCTGGGAACATGTGCAGCAGGCTTTGGACAAACACGAAGGGCTGCTATTC 369
         ||   ||||||::::::|||         :::   :::|||:::|||:::
      87 lyLeuGluHisLeuHisGlnLeuGlnArgGluGlyArgGlyValLeuLeu 103

     370 ATCACGCCGCACATCGGCAGCTACGATTTGGGCGGACGCTACATCAGCCA 419
         ::::::   |||:::   :::   ::::::   ||||||   :::   ::
     104 ValSerGlyHisPheMetThrLeuGluIleCysGlyArgLeuLeuCysAs 120

     420 GCAGCTTCCGTTCCCGCTGACCGCCATGTACAAACCGCCGAAAATCAAAG 469
         :   :::      ||||||::::::||||||:::      :::      :
     120 pTyrVal......ProLeuAlaGlyMetTyrArgArgHisArgAsnProV 135

     470 CGATAGACAAAATCATGCAGGCGGGCAGGGTTCGCGGCAAAGGAAAAACC 519
         ::::::::      ::::::   ||||||:::|||         |||
     135 alLeuGluTrpAlaIleLysArgGlyArgLeuArg......TyrAlaThr 149

     520 GCGCCTACCAGCATACAAGGGGTCAAACAAATCATCAAAGCCCTGCGTTC 569
                 :::   :::   ::::::      ||||||   |||:::
     150 GlnMetPheAlaAsnGluGluLeuArgAlaThrIleLysHisLeuLysAr 166

     570 GGGCGAAGCAACCATCGTCCTGCCCGACCACGTCCCCTCCCCTCAAGAAG 619
         |||                |||||||         :::   ::::::|
     166 gGlyGlyPheLeuTrpTyrAlaProAsp.........GlnAspMetArgG 180

     620 GCGGGGAAGGCGTATGGGTGGATTTCTTCGGCAAACCTGCCTATACCATG 669
         ||   |||   |||::::::   |||||||||   ||||||   |||:::
     180 lyLysGluThrValPheAlaProPhePheGlyMetProAlaAlaThrIle 196

     670 ACGCTGGCGGCAAAATTGGCACACGTCAAAGGCGTGAAAACCCTG...TT 716
         |||   :::   :::|||||||::::::   |||      ::::::   ::
     197 ThrAlaThrHisGlnLeuAlaArgLeuThrGlyCysAlaValValProTy 213

     717 TTTCTGCTGCGAACGCCTGCCTGGCGGACAAGGTTTCGATTTGCACATCC 766
         :|||      :::|||      ||||||   ::::::   |||   |||
     213 rPhe......HisArgArgGlnGlyGly...AsnTyrValLeuLysIleA 227

     767 GCCCCGTCCAAGGGGAATTG...AACGGCGACAAAGCCCATGATGCCGCC 813
           |||      :::::::::   :::   |||   |||   |||:::|||
     227 laProProLeuAlaAspPheProSerGluAspLeuAlaAlaAspThrAla 243

     814 GTGTTCAACCGCAATGCCGAATATTGGATACGCCGTTTTCCGACGCAGTA 863
         :::|||:::      |||      |||   :::   |||   |||||
     244 ArgIleAsnGlnValIleGluThrMetIleAspGluAlaProAspGlnTy 260

     864 TCTGTTTATGTACAACCGCTACAAA 888
         ||||:::::::::::::|||:::|||
     260 rLeuTrpIleHisArgArgPheLys 268

seq_name: sp_bacteria:Q9API8

seq_documentation_block:
ID   Q9API8      PRELIMINARY;      PRT;   292 AA.
AC   Q9API8;
DT   01-JUN-2001 (TrEMBLrel. 17, Created)
DT   01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   PUTATIVE LAYROYL ACYLTRANSFERASE.
OS   Methylobacterium extorquens.
OC   Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC   Methylobacterium group; Methylobacterium.
OX   NCBI_TaxID=408;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=AM1;
RX   MEDLINE=21142531; PubMed=11208803;
RA   Korotkova N., Lidstrom M.E.;
RT   "Connection between Poly-beta-Hydroxybutyrate Biosynthesis and Growth
RT   on C1 and C2 Compounds in the Methylotroph Methylobacterium extorquens
RT   AM1.";
RL   J. Bacteriol. 183:1038-1046(2001).
DR   EMBL; AF287907; AAK11538.1; -.
KW   Transferase; Acyltransferase.
SQ   SEQUENCE   292 AA;  32214 MW;  02AD260FA71E7C11 CRC64;


alignment_scores:
           Quality:  172.50              Length:    294
             Ratio:  1.150                Gaps:     12
 Percent Similarity:  51.020   Percent Identity: 24.830

alignment_block:
 US-09-303-518D-569 x Q9API8   ..

 Align seg 1/1  to: Q9API8  from: 1  to: 292

     102 TCTGCACACGCTGGGAAACCGGCTCGGACATCTGGCGTTTTACCTTTTAA 151
         ||||||   :::|||:::   :::|||      |||      ||||||
       8 SerAlaAlaGlyGlyArgLeuValArgLeuIleGly......ProPheLe 22

     152 AGGAAGACCGCGCGCGCATCGTCGCCAATATGCGTCAGGCAGGCATGAAT 201
                 |||:::    :::   |||:::|||   |||      :::
      22 uProSerHisArgThr.AlaMetAspAsnLeuArgAlaAlaTyrProLys 38

     202 CCCGACCCCAAAACGGTCAAAGCCGTTTTTGCGGAAACG.......... 240
         |||   :::   ::::::|||:::      |||:::
      39 GluAspGluArgArgLeuArgAlaIleAlaLeuGluAlaTrpAspAsnLe 55

     241 .GCAAAAGGCGGTTTGGAACTTGCC.........CCCGCGTTTTTCAGAA 280
```



```
seq_name: sp_bacteria:Q9L7U9

seq_documentation_block:
ID   Q9L7U9      PRELIMINARY;      PRT;   318 AA.
AC   Q9L7U9;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
DE   LIPID A ACYLTRANSFERASE.
GN   MSBB.
OS   Haemophilus ducreyi.
OC   Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
OC   Haemophilus.
OX   NCBI_TaxID=730;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Munson R.S. Jr., Bozue J.A.;
RT   "Identification of the msbB gene in Haemophilus ducreyi.";
RL   Submitted (DEC-1999) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF219260; AAF33777.1; -.
KW   Transferase; Acyltransferase.
SQ   SEQUENCE   318 AA;  36799 MW;  E4C1F82FB88BE37A CRC64;


alignment_scores:
            Quality:  172.50              Length:    262
              Ratio:  1.113                 Gaps:     10
 Percent Similarity:  59.160   Percent Identity: 24.427

alignment_block:
 US-09-303-518D-569 x Q9L7U9   ..

 Align seg 1/1  to: Q9L7U9  from: 1  to: 318
```



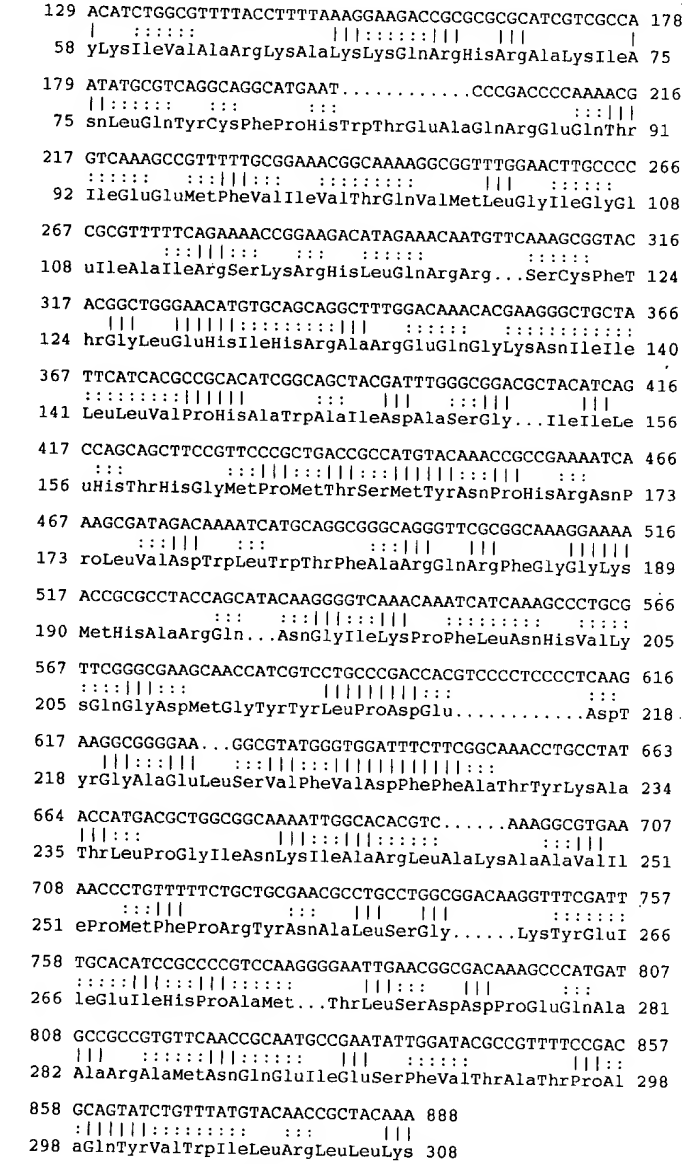
```
      263 lnMetGluIleLeuProAlaIle...GluPheSerGlySerSerAlaGln 278

      805 GATGCCGCCGTGTTCAACCGCAATGCCGAATATTGGATACGCCGTTTTCC 854
      279 SerAlaArgGluMetAsnLysValIleGluTyrPheValHisLysAsnPr 295

      855 GACGCAGTATCTGTTTATGTACAACCGCTACAAA 888
      295 oAlaGlnTyrValTrpIleLeuArgLeuLeuLys 306

seq_name: sp_bacteriap:O84013

seq_documentation_block:
ID   O84013      PRELIMINARY;      PRT;   455 AA.
AC   O84013;
DT   01-NOV-1998 (TrEMBLrel. 08, Created)
DT   01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   ACYLTRANSFERASE.
GN   HTRB OR CT010.
OS   Chlamydia trachomatis.
OC   Bacteria; Chlamydiales; Chlamydiaceae; Chlamydia.
OX   NCBI_TaxID=813;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=D/UW-3/CX;
RX   MEDLINE=99000809; PubMed=9784136;
RA   Stephens R.S., Kalman S., Lammel C.J., Fan J., Marathe R., Aravind L.,
RA   Mitchell W.P., Olinger L., Tatusov R.L., Zhao Q., Koonin E.V.,
RA   Davis R.W.;
RT   "Genome sequence of an obligate intracellular pathogen of humans:
RT   Chlamydia trachomatis.";
RL   Science 282:754-759(1998).
DR   EMBL; AE001275; AAC67600.1; -.
KW   Transferase; Acyltransferase; Complete proteome.
SQ   SEQUENCE   455 AA;  52058 MW;  0404B6326C67ACCF CRC64;


alignment_scores:
           Quality:  172.50            Length:    342
             Ratio:  1.033               Gaps:     12
 Percent Similarity:  48.830   Percent Identity: 23.977

alignment_block:
 US-09-303-518D-569 x O84013   ..

 Align seg 1/1  to: O84013  from: 1  to: 455

       22 CTGTTTCCCCCTTTGCGAACCGCCATGCACATCCTGTTGACCGCCCTG.. 69
        2 LeuPheLysArgLeuArgThrGlyGlyLysIleLeuValAspHisLeuVa 18

       70 .....CTCAAATGCCTCTCCCTGCTGCCGCTTT 97
       18 lTyrGlyLeuGlyLeuGlyValLeuThrIleLeuArgLeuLeuProArgS 35

       98 CCTGTCTGCACACGCTGGGAAACCGGCTCGGACATCTGGCGTTTTACCTT 147
       35 erSerLeuArgLeuPheSerLysGlyLeuGlyThrAlaLeuPheTyrPhe 51

      148 TTAAAGGAAGACCGCGCGCGCATCGTCGCCAATATG..... 183
       52 IleSerAspValArgLysThrAlaLeuThrAsnLeuAlaLeuAlaPhePr 68

      184 .....CGTCAGGCA..... 192
       68 oGluLysSerPheAlaGluArgTyrGlnIleAlaArgGlnSerValGlnG 85

      193 .....GGCATGAATCCCGACCCCAAAACGGTCAAA 222
       85 lnMetIleIleThrPheValGluLeuAlaThrValAspLysPheAlaLys 101

      223 GCCGTTTTTGCGGAAACGGCAAAAGGCGGTTTGGAACTTGCCCCC...GC 269
      102 HisIleAspGluMetIleAlaIleAlaThrSerGluAspAlaProGluGl 118

      270 GTTTTTCAGAAAACCGGAAGACATAGAAACAATGTTCAAAGCGGTACACG 319
      118 yPhePhe......ProGluGluValSerSerGlnGlnGluLeu..... 130

      320 GCTGGGAACATGTGCAGCAGGCTTTGGACAAACACGAAGGGCTGCTATTC 369
      131 .....AspHisPhePheSerArgLeuAspArgGlnGluGlyAlaIleLeu 145

      370 ATCACGCCGCACATCGGCAGCTACGATTTGGGCGGACGCTACATCAGCCA 419
      146 PheCysGlyHisGlnAlaAsnTrpGluLeuProPheLeuTyrIleThrLy 162

      420 GCAGCTTCCGTTCCCGCTGACCGCCATGTACAAACCGCCGAAAATCAAAG 469
      162 sArg......TyrProGlyLeuAlaPheAlaLysProValLysAsnArgA 177

      470 CGATAGACAAAATCATGCAGGCGGGCAGGGTTCGCGGCAAAGGAAAAACC 519
      177 rgLeuAsnGlnLysIleIleSerLeuArgGluSerPheGlnGlyLysIle 193

      520 GCGCCTACCAGCATACAAGGGGTCAAACAAATCATCAAAGCCCTGCGTTC 569
      194 ValProProGln...AsnAlaIleAsnGlnAlaLeuArgAlaLeuHisAr 209

      570 GGGCGAAGCAACCATCGTCCTGCCCGACCACGTCCCCTCCCCTCAAGAAG 619
      209 gGlyGluValValGlyIleValGlyAspGlnValLeuLeuSerSerGluT 226

      620 GCGGGGAAGGCGTATGGGTGGATTTCTTCGGCAAACCTGCCTATACCATG 669
      226 yrSer.....TyrProLeuPheGlySerGlnAlaPheThrThr 238

      670 ACGCTGGCGGCAAAATTGGCACACGTCAAAGGCGTGAAAACCCTGTTTTT 719
      239 ThrSerProAlaLeuLeuAlaTyrLysThrLysLysThrValIleAlaVa 255

      720 CTGCTGCGAACGCCTGCCTGGCGGACAAGGTTTCGATTTGCACATCCGCC 769
      255 lAlaIleTyrArgLysProAsnGlyAsnTyrLeuValVal..... 268

      770 CCGTCCAAGGGGAATTGAACGGCGACAAAGCCCATGATGCCGCCGTGTTC 819
      269 .....ProSerLysAlaPhe 273

      820 AACCGCAATGCCGAATATTGGATACGC..... 846
      274 HisAlaAsnThrGluLeuSerIleArgGluSerThrGluGlnLeuMetAs 290

      847 .....CGTTTT.....CCGACGCAGT 862
      290 pArgLeuMetArgPheLeuGluLysGlyIleThrCysLysProGluGlnT 307

      863 ATCTGTTTATGTACAACCGCTACAAA 888
      307 rpLeuTrpLeuHisLysArgTrpLys 315

seq_name: sp_bacteriap:Q9ZCL1

seq_documentation_block:
ID   Q9ZCL1      PRELIMINARY;      PRT;   290 AA.
AC   Q9ZCL1;
DT   01-MAY-1999 (TrEMBLrel. 10, Created)
DT   01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   LIPID A BIOSYNTHESIS LAUROYL ACYLTRANSFERASE (HTRB).
GN   RP718.
OS   Rickettsia prowazekii.
OC   Bacteria; Proteobacteria; alpha subdivision; Rickettsiales;
OC   Rickettsiaceae; Rickettsieae; Rickettsia.
```

```
OX   NCBI_TaxID=782;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=MADRID E;
RX   MEDLINE=99039499; PubMed=9823893;
RA   Andersson S.G.E., Zomorodipour A., Andersson J.O.,
RA   Sicheritz-Ponten T., Alsmark U.C.M., Podowski R.M., Naeslund A.K.,
RA   Eriksson A.-S., Winkler H.H., Kurland C.G.;
RT   "The genome sequence of Rickettsia prowazekii and the origin of
RT   mitochondria.";
RL   Nature 396:133-140(1998).
DR   EMBL; AJ235273; CAA15149.1; -.
KW   Complete proteome.
SQ   SEQUENCE   290 AA;  33809 MW;  991FF50AB841D5B3 CRC64;


alignment_scores:
            Quality:  169.50            Length:    206
              Ratio:  1.274               Gaps:      9
 Percent Similarity:  64.563   Percent Identity: 22.816

alignment_block:
 US-09-303-518D-569 x Q9ZCL1   ..

 Align seg 1/1  to: Q9ZCL1  from: 1  to: 290

     289 GACATAGAAACAATGTTCAAAGCGGTACACGGCTGGGAACATGTGCAGCA 338
         :::::::|||:::   ::::::   :::   |||   |||:::::::::::
      95 GluLeuGluSerArgIleGluIleIle...GlyLysGluAsnIleLysLy 110

     339 GGCTTTGGACAAACACGAAGGGCTGCTATTCATCACGCCGCACATCGGCA 388
         :      |||      :::   :::|||:::::::::   |||:::::::
     110 sLeuGluAsp......GlnProPheLeuLeuPheSerGlyHisPheAlaA 125

     389 GCTACGATTTGGGCGGACGCTACATCAGCCAGCAGCTTCCGTTCCCGCTG 438
         :::::|||::::::   :::   :::::::::      |||      :::
     125 snTrpAspIleSerLeuLysValLeuLysLysTyrTyrPro...LysVal 140

     439 ACCGCCATGTACAAACCGCCGAAAATCAAAGCGATAGACAAAATCATGCA 488
         :::::::::|||:::      :::        ::::::|||::::::::
     141 AlaValIleTyrArgLysAlaAsnAsnProTyrValAsnLysLeuValAs 157

     489 GGCGGGCAGGGTTCGCGGCAAAGGAAAAACCGCGCCTACCAGCATACAAG 538
         :   :::|||:::      |||   :::      |||   :::   :::|
     157 nGluSerArgAlaGlyAspLysLeuArgLeuIleProLysGlySerGluG 174

     539 GGGTCAAACAAATCATCAAAGCCCTGCGTTCGGGCGAAGCAACCATCGTC 588
         ||:::::::::::::::::::|||:::::::::|||||:::   ::::::
     174 lyIleArgSerLeuValArgAlaIleLysGluGlyGluSerIleValMet 190

     589 CTGCCCGACCACGTCCCCTCCCCTCAAGAAGGCGGGGAAGGCGTATGGGT 638
         |||   |||                  |||:::   ::::::|||:::   ||
     191 LeuValAsp..............GlnLysMetAsnAspGlyIleGluVa 202

     639 GGATTTCTTCGGCAAACCTGCCTATACCATGACGCTGGCGGCAAAATTGG 688
         |   |||:::|||:::|||:::   |||   :::      ||||||:::|
     202 lProPheLeuGlyArgProSerMetThrAlaSerSerIleAlaLysIleA 219

     689 CACACGTCAAAGGCGTGAAAACCCTGTTTTTCTGCTGCGAACGCCTG... 735
         ||               |||      ::::::   |||:::   :::
     219 la......LeuGlnTyrLysTyrProIleIleProCysGlnIleIleArg 233

     736 CCTGGCGGACAAGGTTTCGATTTGCACATCCGCCCC...GTCCAAGGGGA 782
             |||:::   |||   :::   ::::::|||   ::::::   ||
     234 ThrLysGlySerTyrPheLysValIleValHisProGlnLeuLysPheGl 250

     783 ATTGAACGGCGACAAAGCCCATGATGCC......GCCGTGTTCAACCGCA 826
         |   :::|||||||:::      |||:::           :::||||||   |
     250 uGlnThrGlyAspAsnLysValAspCysTyrAsnIleMetPheAsnIleA 267

     827 AT......GCCGAATATTGGATACGCCGTTTTCCGACGCAGTATCTGTTT 870
         ||               |||::::::::::   |||:::|||:::::::::
     267 snGlnIleLeuGlyGluTrpValLysGlnAsnProAlaGlnTrpPheTrp 283

     871 ATGTACAACCGCTACAAA 888
         ::::::|||||||:::|||
     284 PheHisAsnArgTrpLys 289

seq_name: sp_bacteria:Q9AKA6

seq_documentation_block:
ID   Q9AKA6      PRELIMINARY;      PRT;   290 AA.
AC   Q9AKA6;
DT   01-JUN-2001 (TrEMBLrel. 17, Created)
DT   01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   HTRB PROTEIN.
GN   HTRB.
OS   Rickettsia typhi.
OC   Bacteria; Proteobacteria; alpha subdivision; Rickettsiales;
OC   Rickettsiaceae; Rickettsieae; Rickettsia.
OX   NCBI_TaxID=785;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=WILMINGTON;
RX   MEDLINE=21219194; PubMed=11319266;
RA   Andersson J.O., Andersson S.G.E.;
RT   "Pseudogenes, junk DNA and the dynamics of Rickettsia genomes.";
RL   Mol. Biol. Evol. 18:829-839(2001).
DR   EMBL; AJ293328; CAC33762.1; -.
SQ   SEQUENCE   290 AA;  33814 MW;  2E9B2D3193B70EDD CRC64;


alignment_scores:
            Quality:  166.00            Length:    196
              Ratio:  1.339               Gaps:      7
 Percent Similarity:  63.265   Percent Identity: 22.959

alignment_block:
 US-09-303-518D-569 x Q9AKA6   ..

 Align seg 1/1  to: Q9AKA6  from: 1  to: 290

     319 GGCTGGGAACATGTGCAGCAGGCTTTGGACAAACACGAAGGGCTGCTATT 368
         |||   |||:::::::|||:::      |||      :::   :::|||::
     104 GlyLysGluAsnIleGlnLysLeuGluAsp......GlnProPheLeuLe 118

     369 CATCACGCCGCACATCGGCAGCTACGATTTGGGCGGACGCTACATCAGCC 418
         :::::::   |||:::::::::::|||::::::   :::   :::::::
     118 uPheSerGlyHisPheAlaAsnTrpAspIleSerLeuLysValLeuLysL 135

     419 AGCAGCTTCCGTTCCCGCTGACCGCCATGTACAAACCGCCGAAAATCAAA 468
         ::      |||      :::::::::::::|||:::      :::
     135 ysTyrTyrPro...LysValAlaValIleTyrArgLysAlaAsnAsnPro 150

     469 GCGATAGACAAAATCATGCAGGCGGGCAGGGTTCGCGGCAAAGGAAAAAC 518
         ::::::|||:::::::::   :::|||:::      |||   :::
     151 TyrValAsnLysLeuValAsnGluSerArgAlaGlyAspLysLeuArgLe 167

     519 CGCGCCTACCAGCATACAAGGGGTCAAACAAATCATCAAAGCCCTGCGTT 568
           |||   :::   :::|||::::::   :::::::::|||:::::::
     167 uIleProLysGlySerGluGlyIleArgAlaLeuValSerAlaIleLysG 184

     569 CGGGCGAAGCAACCATCGTCCTGCCCGACCACGTCCCCTCCCCTCAAGAA 618
         ::|||||||:::   ::::::|||   |||               |||:::
     184 luGlyGluSerIleValMetLeuValAsp.............GlnLys 195

     619 GGCGGGGAAGGCGTATGGGTGGATTTCTTCGGCAAACCTGCCTATACCAT 668
         ::::::|||:::   |||   |||:::|||:::|||:::   |||
     196 MetAsnAspGlyIleGluValProPheLeuGlyArgProSerMetThrAl 212

     669 GACGCTGGCGGCAAAATTGGCACACGTCAAAGGCGTGAAAACCCTGTTTT 718
         :::      ||||||:::|||               |||      ::::
     212 aSerAlaIleAlaLysIleAla......LeuGlnTyrLysTyrProIleI 227

     719 TCTGCTGCGAACGCCTG...CCTGGCGGACAAGGTTTCGATTTGCACATC 765
```

```
      ::   |||:::   :::         |||:::   |||   :::   :::
    227 leProCysGlnIleIleArgThrLysGlySerTyrPheLysValIleVal 243

    766 CGCCCC...GTCCAAGGGGAATTGAACGGCGACAAAGCCCATGATGCC.. 810
        :::|||   ::::::   |||   :::|||||||:::      |||:::
    244 HisProGlnLeuLysPheGluGlnThrGlyAspAsnLysValAspCysTy 260

    811 .........GCCGTGTTCAACCGCAATGCCGAATATTGGATACGCCGTT 850
                  :::|||:::            |||:::::::::
    260 rAsnIleMetPheTyrIleAsnGlnIleLeuGlyGluTrpValLysGlnA 277

    851 TTCCGACGCAGTATCTGTTTATGTACAACCGCTACAAA 888
        |||:::|||:::::::::::::::|||||||:::|||
    277 snProAlaGlnTrpPheTrpPheHisAsnArgTrpLys 289

seq_name: sp_bacteriap:Q9PL30

seq_documentation_block:
ID   Q9PL30      PRELIMINARY;      PRT;   453 AA.
AC   Q9PL30;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   LIPID A BIOSYNTHESIS LAUROYL ACYLTRANSFERASE, PUTATIVE.
GN   TC0278.
OS   Chlamydia muridarum.
OC   Bacteria; Chlamydiales; Chlamydiaceae; Chlamydia.
OX   NCBI_TaxID=83560;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=MOPN / NIGG;
RX   MEDLINE=20150255; PubMed=10684935;
RA   Read T.D., Brunham R.C., Shen C., Gill S.R., Heidelberg J.F.,
RA   White O., Hickey E.K., Peterson J., Utterback T., Berry K., Bass S.,
RA   Linher K., Weidman J., Khouri H., Craven B., Bowman C., Dodson R.,
RA   Gwinn M., Nelson W., DeBoy R., Kolonay J., McClarty G., Salzberg S.L.,
RA   Eisen J., Fraser C.M.;
RT   "Genome sequences of Chlamydia trachomatis MoPn and Chlamydia
RT   pneumoniae AR39.";
RL   Nucleic Acids Res. 28:1397-1406(2000).
DR   EMBL; AE002295; AAF39146.1; -.
DR   TIGR; TC0278; -.
KW   Complete proteome.
SQ   SEQUENCE   453 AA;  51463 MW;  6221515A00A093FF CRC64;


alignment_scores:
           Quality:  164.50            Length:    336
             Ratio:  0.962              Gaps:     14
 Percent Similarity:  50.893   Percent Identity: 24.702

alignment_block:
 US-09-303-518D-569 x Q9PL30   ..

 Align seg 1/1  to: Q9PL30  from: 1  to: 453

     22 CTGTTTCCCCCTTTGCGAACCGCCATGCACATCCTGTTG.......... 60
        ||||||      ||||||::::::      :::|||:::
      2 LeuPheLysMetLeuArgSerGlyGlyLysValLeuValAspHisLeuVa 18

     61 ............ACCGCCCTGCTCAAATGCCTCTCCCTGCTGCCGCTTT 97
                    :::::::|||      |||   |||||||||   |
     18 lTyrGlyLeuGlyLeuGlyIleLeuThrIleLeuArgLeuLeuProArgS 35

     98 CCTGTCTGCACACGCTGGGAAACCGGCTCGGACATCTGGCGTTTTACCTT 147
        ||   |||:::   :::|||:::   ||||||         ||||||:::
     35 erSerLeuGlnLeuPheGlyLysGlyLeuGlyThrThrIlePheTyrVal 51

    148 TTAAAGGAAGACCGCGCGCGCATCGTCGCCAATATGCGTCAGGCAGGCAT 197
        :::::::::   |||         ::::::|||:::      |||
     52 IleSerAspPheArgLysThrAlaLeuThrAsnLeuAlaLeuAla..... 66

    198 GAATCCCGACCCC.............................AAAACGG 217
```

```
        |||:::                                  |||:::|
     67 .PheProGluLysSerPheThrGluArgTyrGlnIleAlaLeuLysSerV 83

    218 TCAAAGCCGTT.........TTTGCGGAAACGGCAAAA............ 246
        ||:::   |||         |||   |||   |||
     83 alGlnGlnValIleIleThrPheIleGluLeuAlaThrValAspLysPhe 99

    247 ............................GGCGGTTTGGAACTTGCCCC 266
                                    ::::::   |||   |||||
    100 AlaLysHisIleAspGluIleIleThrIleAlaSerSerGluAspAlaPr 116

    267 C...GCGTTTTTCAGAAAACCGGAAGACATAGAAACAATGTTCAAAGCGG 313
        |   :::|||||      ||||||:::::::::::::::   :::
    116 oGluGlyPhePhe......ProGluGluValSerSerGlnGlnGluLeu. 130

    314 TACACGGCTGGGAACATGTGCAGCAGGCTTTGGACAAACACGAAGGGCTG 363
                  ::::::      :::   ||||||::::::||||||
    131 ..........AsnAsnPhePheSerArgLeuAspGlnGlnGluGlyAla 143

    364 CTATTCATCACGCCGCACATCGGCAGCTACGATTTGGGCGGACGCTACAT 413
        :::::::::      |||   ::::::::::|||          |||||
    144 IleLeuPheCysGlyHisGlnAlaAsnTrpGluLeuProPheLeuTyrIl 160

    414 CAGCCAGCAGCTTCCGTTCCCGCTGACCGCCATGTACAAACCGCCGAAAA 463
        |:::::::::      :::|||      |||:::   ||||||   |||
    160 eThrLysArg......TyrProGlyLeuAlaPheAlaLysProValLysA 175

    464 TCAAAGCGATAGACAAAATCATGCAGGCGGGCAGGGTTCGCGGCAAAGGA 513
                ::::::::   :::   :::   |||          :::|||
    175 snProArgLeuAsnArgLysIleIleSerLeuArgGluSerPheGlnGly 191

    514 AAAACCGCGCCTACCAGCATACAAGGGGTCAAACAAATCATCAAAGCCCT 563
        |||   :::|||   :::   ::::::::::::|||   ::::::|||||
    192 LysIleValProProGln...AsnAlaIleAsnGlnAlaLeuArgAlaLe 207

    564 GCGTTCGGGCGAAGCAACCATCGTCCTGCCCGACCACGTCCCCTCCCCTC 613
        |::::::||||||::::::   ::::::   |||:::|||        :
    207 uHisLysGlyGluValValGlyIleValGlyAspGlnValLeuLeuSerS 224

    614 AAGAAGGCGGGGAAGGCGTATGGGTGGATTTCTTCGGCAAACCTGCCTAT 663
        :::::   :::                ::::::::::   |||:::
    224 erGlnTyrSer............TyrProLeuPheGlySerGlnAlaPhe 236

    664 ACCATGACGCTGGCGGCAAAATTGGCACACGTCAAAGGCGTGAAAACCCT 713
        |||   |||      |||   ||||||:::                :::::
    237 ThrThrThrSerProAlaLeuLeuAlaTyrLysThrArgLysProValIl 253

    714 GTTTTTCTGCTGCGAACGCCTGCCTGGCGGACAA.............. 747
        :      :::      |||   |||:::|||:::
    253 eAlaValAlaIleTyrArgGlnProAsnGlyAsnTyrLeuValValProS 270

    748 ................GGTTTCGATTTGCACATCCGCCCCGTCCAAGGG 780
                        :::   :::|||   ||||||
    270 erLysAlaPheTyrAlaAsnThrGluLeuSerIleArgGluSerThrGlu 286

    781 GAATTGAACGGCGACAAAGCCCATGATGCCGCCGTGTTCAACCGCAATGC 830
        :::|||      ||||||                    :::   |||
    287 GlnLeuMet...AspLys.................LeuMetArgPheLe 296

    831 CGAATATTGGATACGCCGTTTTCCGACGCAGTATCTGTTTATGTACAACC 880
        |||      |||         |||   |||:::|||:::::::::::::|
    296 uGluLysGlyIleAlaCysLysProGlnTrpLeuTrpLeuHisLysA 313

    881 GCTACAAA 888
        ||:::|||
    313 rgTrpLys 315

seq_name: sp_bacteriap:Q9CME5

seq_documentation_block:
ID   Q9CME5      PRELIMINARY;      PRT;   313 AA.
AC   Q9CME5;
DT   01-JUN-2001 (TrEMBLrel. 17, Created)
```

```
DT   01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   HYPOTHETICAL PROTEIN PM0885.
GN   PM0885.
OS   Pasteurella multocida.
OC   Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
OC   Pasteurella.
OX   NCBI_TaxID=747;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=PM70;
RX   MEDLINE=21145866; PubMed=11248100;
RA   May B.J., Zhang Q., Li L.L., Paustian M.L., Whittam T.S., Kapur V.;
RT   "Complete genomic sequence of Pasteurella multocida Pm70.";
RL   Proc. Natl. Acad. Sci. U.S.A. 98:3460-3465(2001).
DR   EMBL; AE006127; AAK02969.1; -.
KW   Hypothetical protein; Complete proteome.
SQ   SEQUENCE   313 AA;  35956 MW;  23456DB87C45E535 CRC64;


alignment_scores:
            Quality:  163.50              Length:    302
              Ratio:  1.062                 Gaps:     12
 Percent Similarity:  50.993   Percent Identity: 25.166

alignment_block:
 US-09-303-518D-569 x Q9CME5   ..

 Align seg 1/1  to: Q9CME5  from: 1  to: 313

      64 GCCCTGCTCAAATGCCTCTCCCTGCTGCCGCTTTCCTGTCTGCACACGCT 113
         ||||||   :::   :::      ||||||         ||||||   ||
      26 AlaLeuTrpArgLeuIleValCysLeuProTyrProIleLeuHisLysLe 42

     114 GGGAAACCGGCTCGGACAT...CTGGCGTTTTACCTTTTAAAGGAAGACC 160
         |||||||   ||||||      ||||||         :::   :::   |
      42 uGlyAsnGlyLeuGlyTrpLeuLeuAlaLysSerThrValGlyGlnArgA 59

     161 GCGCGCGCATCGTC..................................... 174
         |||||:::|||:::
      59 rgAlaAsnIleAlaArgArgAsnLeuGlnLeuCysPheProHisTyrSer 75

     175 ..........................GCCAATATGCGTCAGGCAGGCAT 197
                                    ||||||:::::   :::|||||
      76 AlaAspGlnIleGluThrLeuLeuHisAlaAsnLeuLysAlaThrGlyMe 92

     198 GAATCCCGACCCCAAAACGGTCAAAGCCGTTTTTGCGGAAACGGCAAAAG 247
         |                        |||:::                 :
      92 t.......................AlaIle..................... 94

     248 GCGGTTTGGAACTTGCCCCCGCGTTTTTCAGAAAACCGGAAGACATAGAA 297
           :::|||   :::   |||:::|||   :::         |||:::
      95 .....IleGluThrGlyMetAlaTrpPheTrpSerAspAlaArgIleLys 109

     298 ACAATGTTCAAAGCGGTACACGGCTGGGAACATGTGCAGCAGGCTTTGGA 347
             |||   ::::::|||   |||:::::::::::::
     110 LysTrpSerLys...IleGluGlyLeuGluTyrLeuLysGluTyr..... 123

     348 CAAACACGAAGGGCTGCTATTCATCACGCCGCACATCGGCAGCTACGATT 397
          |||::::::|||:::::::::::::     |||:::   :::   :::|
     124 .LysGlnAspGlyValIleLeuValGlyValHisPheLeuThrLeuGluL 140

     398 TGGGCGGACGCTACATCAGCCAGCAGCTTCCGTTCCCGCTGACCGCCATG 447
         |||||:::|||   ::::::      |||      |||      ::::::
     140 euGlyAlaArgIleValGly......LeuHisGlnProGlyIleGlyVal 154

     448 TACAAACCGCCGAAAATCAAAGCGATAGACAAAATCATGCAGGCGGGCAG 497
         |||:::|||                :::|||   ::::::      |||||
     155 TyrArgProAsnAspAsnProLeuMetAspTrpLeuGlnValGlnGlyAr 171

     498 GGTTCGCGGCAAAGGAAAAACCGCGCCTACCAGCATACAAGGGGTCAAAC 547
         |:::|||::::::                    :::   ::::::
     171 gLeuArgSerAsnLysAspLeuLeuAspArg......LysAspLeuArgG 186

     548 AAATCATCAAAGCCCTGCGTTCGGGCGAAGCAACCATCGTCCTGCCCGAC 597
           :::|||:::|||||||||:::||||||                |||||
     186 lyMetIleArgAlaLeuArgGlnGlyGluIleIleTrpTyrAlaProAsp 202

     598 CACGTCCCCTCCCCTCAAGAAGGCGGG......GAAGGCGTATGGGTGGA 641
         |||                :::   |||      ::::::|||:::|||
     203 His..............AspTyrGlyArgLysAsnAlaValPheValPr 214

     642 TTTCTTCGGCAAACCTGCCTATACCATGACGCTGGCGGCA......AAAT 685
          ||||||:::   |||      :::   |||   ::::::        |
     214 oPhePheAlaValProAspAlaAlaThrThrThrGlySerTyrTyrLeuL 231

     686 TGGCACACGTCAAAGGCGTGAAAACCCTGTTTTTCTGCTGCGAACGCCTG 735
         ||                  |||::::::   |||:::      |||
     231 euLysSerSerProThrSerLysValIleProPheAlaProLeuArgAsn 247

     736 CCTGGCGGACAAGGTTTCGATTTGCACATCCGC...CCCGTCCAAGGGGA 782
           |||:::|||:::   :::   |||      ||||||:::   ::
     248 SerAspGlySerGlyTyrThrValLysIleSerProProValAspPheSe 264

     783 ATTGAACGGCGACAAAGCCCATGATGCCGCCGTGTTCAACCGCAATGCCG 832
         :        |||:::::::::   |||:::   :::|||:::   :::|
     264 rAspIleThrAspGluThrArgIleAlaThrArgMetAsnGlnValValG 281

     833 AATATTGGATACGCCGTTTTCCGACGCAGTATCTGTTTATGTACAACCGC 882
         ||      |||   :::         ||||||:::::::::::::::|||
     281 luThrGluIleMetLysGlyValGluGlnTyrMetTrpLeuHisArgArg 297

     883 TACAAA 888
         :::|||
     298 PheLys 299

seq_name: sp_bacteriap:Q9CNR4

seq_documentation_block:
ID   Q9CNR4      PRELIMINARY;      PRT;   318 AA.
AC   Q9CNR4;
DT   01-JUN-2001 (TrEMBLrel. 17, Created)
DT   01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   MSBB.
GN   MSBB OR PM0362.
OS   Pasteurella multocida.
OC   Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
OC   Pasteurella.
OX   NCBI_TaxID=747;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=PM70;
RX   MEDLINE=21145866; PubMed=11248100;
RA   May B.J., Zhang Q., Li L.L., Paustian M.L., Whittam T.S., Kapur V.;
RT   "Complete genomic sequence of Pasteurella multocida Pm70.";
RL   Proc. Natl. Acad. Sci. U.S.A. 98:3460-3465(2001).
DR   EMBL; AE006072; AAK02446.1; -.
KW   Complete proteome.
SQ   SEQUENCE   318 AA;  36768 MW;  F91DDAB6D9C46375 CRC64;


alignment_scores:
            Quality:  162.50              Length:    277
              Ratio:  1.016                 Gaps:      9
 Percent Similarity:  57.762   Percent Identity: 22.383

alignment_block:
 US-09-303-518D-569 x Q9CNR4   ..

 Align seg 1/1  to: Q9CNR4  from: 1  to: 318

      79 CTCTCCCTGCTGCCGCTTTCCTGTCTGCACACGCTGGGAAACCGGCTCGG 128
         |||:::::::::|||:::              |||::::::   :::||
      42 LeuAlaPheValProPheArgLeuArgAspLysLeuAlaSerTyrIleGl 58
```



```
seq_name: sp_bacteria:Q9L7Q3

seq_documentation_block:
ID   Q9L7Q3      PRELIMINARY;      PRT;   317 AA.
AC   Q9L7Q3;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
DE   HTRB.
GN   HTRB.
OS   Haemophilus ducreyi.
OC   Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
OC   Haemophilus.
OX   NCBI_TaxID=730;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Munson R.S. Jr., Bozue J.A.;
RT   "Identification of the htrB gene in Haemophilus ducreyi.";
RL   Submitted (JAN-2000) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF221008; AAF34642.1; -.
SQ   SEQUENCE   317 AA;  36521 MW;  8359FE1D5945612B CRC64;


alignment_scores:
           Quality:  161.50              Length:    286
             Ratio:  1.003                Gaps:     12
 Percent Similarity:  56.294   Percent Identity: 25.874

alignment_block:
 US-09-303-518D-569 x Q9L7Q3   ..

 Align seg 1/1  to: Q9L7Q3  from: 1  to: 317
```

```
      ||||||:::                    |||   ::::::   :::
    221 PhePheAlaValGluGlnAlaCysThrThrAlaGlySerHisMetLeuLe 237

    693 CGTCAAAGGCGTGAAAACCCTGTTT......TTCTGCTGCGAACGCCTGC 736
        ::::::   :::::::::          |||         |||
    237 uArgSerAlaProAsnSerIleValValProPheThrProIleArgHisT 254

    737 CTGGCGGACAAGGTTTCGATTTGCACATCCGCCCCGTCCAAGGGGAATTG 786
            :::|||:::   :::   |||   |||:::      ::::::
    254 yrAspPheSerGlyTyrThrValLysIleSerProAlaVal...AspPhe 269

    787 AACGGC......GACAAAGCCCATGATGCCGCCGTGTTCAACCGCAATGC 830
        |||         :::|||   :::   |||   :::::::|||:::   ::
    270 AsnAspCysGlnAsnLysIleGluThrAlaArgLeuMetAsnLysValVa 286

    831 CGAATATTGGATACGCCGTTTTCCGACGCAGTATCTGTTTATGTACAACC 880
        :|||      |||   :::      |||||||||:::::::::::::::|
    286 lGluThrGluIleLeuLysAlaGlnThrGlnTyrMetTrpLeuHisArgA 303

    881 GCTACAAA 888
        ||:::|||
    303 rgPheLys 305

seq_name: sp_bacteria:Q9AKF0

seq_documentation_block:
ID   Q9AKF0      PRELIMINARY;      PRT;   290 AA.
AC   Q9AKF0;
DT   01-JUN-2001 (TrEMBLrel. 17, Created)
DT   01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   HTRB PROTEIN.
GN   HTRB.
OS   Rickettsia rickettsii.
OC   Bacteria; Proteobacteria; alpha subdivision; Rickettsiales;
OC   Rickettsiaceae; Rickettsieae; Rickettsia.
OX   NCBI_TaxID=783;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=84-21C;
RX   MEDLINE=21219194; PubMed=11319266;
RA   Andersson J.O., Andersson S.G.E.;
RT   "Pseudogenes, junk DNA and the dynamics of Rickettsia genomes.";
RL   Mol. Biol. Evol. 18:829-839(2001).
DR   EMBL; AJ293329; CAC33715.1; -.
SQ   SEQUENCE   290 AA;  33550 MW;  9FCC5D73E5CBE89A CRC64;


alignment_scores:
            Quality:  154.00            Length:    201
              Ratio:  1.252               Gaps:      8
 Percent Similarity:  61.194   Percent Identity: 23.881

alignment_block:
 US-09-303-518D-569 x Q9AKF0   ..

 Align seg 1/1  to: Q9AKF0  from: 1  to: 290

    310 GCGGTACACGGCTGGGAACATGTGCAGCAGGCTTTGGACAAACACGAAGG 359
        |||:::   |||   |||:::::::::            |||   |||||
    101 AlaIleIleGlyIleGluAsnIleLys...........LysLeuGluGl 113

    360 G......CTGCTATTCATCACGCCGCACATCGGCAGCTACGATTTGGGCG 403
        |      :::|||:::::::::   |||:::::::::::|||::::::
    113 yGlnProPheLeuLeuPheSerGlyHisPheAlaAsnTrpAspIleSerI 130

    404 GACGCTACATCAGCCAGCAGCTTCCGTTCCCGCTGACCGCCATGTACAAA 453
          :::   :::   :::      |||      :::::::::::::|||:::
    130 leLysIleLeuHisLysPheTyrPro...LysValAlaValIleTyrArg 145

    454 CCGCCGAAAATCAAAGCGATAGACAAAATCATGCAGGCGGGCAGGGTTCG 503
            :::         :::::::|||::::::::   :::|||:::
    146 LysAlaAsnAsnProTyrValAsnLysLeuValAsnGluSerArgAlaGl 162

    504 CGGCAAAGGAAAAACCGCGCCTACCAGCATACAAGGGGTCAAACAAATCA 553
            |||   :::      |||   :::   :::|||   :::   ::::
    162 yAspLysLeuArgLeuIleProLysGlyProGluGlySerArgAlaLeuV 179

    554 TCAAAGCCCTGCGTTCGGGCGAAGCAACCATCGTCCTGCCCGACCACGTC 603
        :::::|||:::::::::::::||||||   ::::::|||   |||
    179 alArgAlaIleLysGluSerGluAlaIleValMetLeuValAsp...... 193

    604 CCCTCCCCTCAAGAAGGCGGGGAAGGCGTATGGGTGGATTTCTTCGGCAA 653
                |||:::   ::::::|||:::   |||   |||:::|||
    194 .........GlnLysMetAsnAspGlyIleGluValProPheLeuGlyHi 207

    654 ACCTGCCTATACCATGACGCTGGCGGCAAAATTGGCACACGTCAAAGGCG 703
        ||||||   |||   :::      ||||||:::|||
    207 sProAlaMetThrAlaSerAlaIleAlaLysIleAla......LeuGlnT 222

    704 TGAAAACCCTGTTTTTCTGCTGCGAACGCCTG...CCTGGCGGACAAGGT 750
        |||      ::::::   |||:::   :::          |||:::
    222 yrLysTyrProIleIleProCysGlnIleIleArgThrLysGlySerTyr 238

    751 TTCGATTTGCACATCCGCCCC...GTCCAAGGGGAATTGAACGGCGACAA 797
        |||   :::   :::::::|||   ::::::   |||   :::|||||::
    239 PheLysValIleValHisProGlnLeuLysLeuGluGlnThrGlyAspAs 255

    798 AGCCCATGATGCC............GCCGTGTTCAACCGCAATGCCGAAT 835
        :      |||:::                  :::|||:::
    255 nLysAlaAspCysTyrAsnIleMetLeuAsnIleAsnGlnMetLeuGlyG 272

    836 ATTGGATACGCCGTTTTCCGACGCAGTATCTGTTTATGTACAACCGCTAC 885
        |||:::::::::   |||:::|||:::::::::::::::::|||||:::
    272 luTrpValLysGlnAsnProAlaGlnTrpPheTrpPheHisAsnArgTrp 288

    886 AAA 888
        |||
    289 Lys 289

seq_name: sp_bacteriap:Q98L49

seq_documentation_block:
ID   Q98L49      PRELIMINARY;      PRT;   328 AA.
AC   Q98L49;
DT   01-OCT-2001 (TrEMBLrel. 18, Created)
DT   01-OCT-2001 (TrEMBLrel. 18, Last sequence update)
DT   01-OCT-2001 (TrEMBLrel. 18, Last annotation update)
DE   LIPID A BIOSYNTHESIS LAUROYL ACYLTRANSFERASE.
GN   MLR1179.
OS   Rhizobium loti (Mesorhizobium loti).
OC   Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC   Phyllobacteriaceae; Mesorhizobium.
OX   NCBI_TaxID=381;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=MAFF303099;
RX   MEDLINE=21082930; PubMed=11214968;
RA   Kaneko T., Nakamura Y., Sato S., Asamizu E., Kato T., Sasamoto S.,
RA   Watanabe A., Idesawa K., Ishikawa A., Kawashima K., Kimura T.,
RA   Kishida Y., Kiyokawa C., Kohara M., Matsumoto M., Matsuno A.,
RA   Mochizuki Y., Nakayama S., Nakazaki N., Shimpo S., Sugimoto M.,
RA   Takeuchi C., Yamada M., Tabata S.;
RT   "Complete genome structure of the nitrogen-fixing symbiotic bacterium
RT   Mesorhizobium loti.";
RL   DNA Res. 7:331-338(2000).
DR   EMBL; AP002996; BAB48614.1; -.
KW   Transferase; Acyltransferase; Complete proteome.
SQ   SEQUENCE   328 AA;  37043 MW;  64A39DC394F5E700 CRC64;


alignment_scores:
            Quality:  153.00            Length:    324
              Ratio:  0.922               Gaps:      12
 Percent Similarity:  51.235   Percent Identity: 22.840
```

```
alignment_block:
US-09-303-518D-569 x Q98L49   ..

Align seg 1/1  to: Q98L49  from: 1  to: 328

     10 TTACAATTCAGGCTGTTTCCCCCTTTGCGAACCGCCATGCACATCCTGTT 59
        |||:::||||||              ||||||          :::   |||::
      9 LeuArgPheArgTyrGlyArgGlnLeuArgGlnLeuAsnTyrTrpLeuVa 25

     60 GACC............GCCCTGCTCAAATGCCTCTCCCTGCTGCCGCTTT 97
        ::::            :::::::::   |||   |||||||||:::
     25 lAlaArgAlaAlaMetIleIleIleSerValLeuArgLeuLeuProValA 42

     98 CCTGTCTGCACACGCTGGGAAACCGGCTCGGACATCTGGCGTTTTACCTT 147
        ::          ::::::::::::|||:::::::::      |||
     42 spSerAlaLeuAsnPheAlaAspArgValAlaArg......PheIleGly 56

    148 TTAAAGGAAGACCGCGCGCGCATC...GTCGCCAATATGCGTCAGGCAGG 194
        |||      |||   ::::::   ::::::|||:::|||:::|||
     57 ProLysValGlyArgHisGlnValAlaIleGlyAsnLeuArgLysAlaTy 73

    195 CATGAATCCCGACCCCAAAACGGTCAAAGCCGTTTTTGCGGAAACGGCAA 244
           :::   ::          ::::::|||:::   ::::::
     73 rProGluLysSerGluAspGluIleGlnAlaIleAlaSerAspMetTrpG 90

    245 AAGGCGGTTTGGAACTTGCCCCCGCGTTTTTC..................... 276
          :::      :::||||||       ::::::
     90 lyAsnMetAlaArgLeuAlaAlaGluTyrIlePheLeuAspAlaLeuPhe 106

    277 ..................AGAAAACCGGAAGACATAGAAACAATGTTCAA 308
                           ||||||      :::|||
    107 AspTyrAspProAlaAlaSerLysProGlyArgValGlu........... 119

    309 AGCGGTACACGGCTGGGAACAT...GTGCAGCAGGCTTTGGACAAACACG 355
           |||   |||   ||||||   |||:::   |||   :::|||:::
    120 ....ValLysGlyIleGluHisPheValGluIleAlaAlaGluLysGlnP 135

    356 AAGGGCTGCTATTCATCACGCCGCACATCGGCAGCTACGATTTGGGCGGA 405
           ::::::|||   |||   |||:::|||:::::::::|||
    135 roHisIleValPhe...ThrGlyHisLeuGlyAsnPheGluLeu...... 148

    406 CGCTACATCAGCCAGCAGCTTCCG.....................TTCCC 434
                        ||||||                       :::
    149 ..................LeuProValAlaAlaAlaThrPheGlyMetAs 159

    435 GCTGACCGCCATGTACAAACCGCCGAAAATCAAAGCGATAGACAAAATCA 484
        :::||||||::::::::::||||||:::          :::       :
    159 nIleThrAlaLeuPheArgProProAsnAsnProTyrLeuAlaAspTyrI 176

    485 TGCAGGCGGGCAGGGTTCGCGGCAAAGGAAAAACCGCGCCTACCAGCATA 534
        ::   :::   |||           |||         |||:::
    176 leLeuSerThrArgArgSerThrMetGlyAlaLeuLeuProSerMetAla 192

    535 CAAGGGGTCAAACAAATCATCAAAGCCCTGCGTTCGGGCGAAGCAACCAT 584
           :::         :::      :::|||:::::::|||
    193 GlyAlaSerPheAlaLeuAlaGlyValLeuGluAsnGlyGlyAsnIleGl 209

    585 CGTCCTGCCCGACCACGTCCCCTCCCCTCAAGAAGGCGGGGAAGGCGTAT 634
        ||||||   |||                |||:::   ::::::|||:::
    209 yValLeuValAsp..............GlnLysPheSerAsnGlyLeuA 221

    635 GGGTGGATTTCTTCGGCAAACCTGCCTATACCATGACGCTGGCGGCAAAA 684
           :::   |||||||||:::|||:::   :::      :::   :::
    221 spThrThrPhePheGlyArgProCysGlnSerAsnArgValLeuGlyThr 237

    685 TTGGCACACGTCAAAGGCGTGAAAACCCTGTTTTTCTGCTGCGAACGCCT 734
        ||||||:::           :::           |||   |||||
    238 LeuAlaArgHisTyrAspCysAspValTyrProAlaArgCysIleArgLe 254

    735 GCCTGGCGGACAAGGTTTCGATTTGCACATCCGC..............C 769
        |||||||::::::   |||   |||:::|||:::             |
    254 uProGlyAsnArg...PheArgLeuGluIleGluAspArgLeuValLeuP 270

    770 CCGTCCAAGGGGAATTGAACGGCGACAAAGCCCATGATGCCGCCGTGTTC 819
           ||      ::::::   :::   |||              :::   ::::::
    270 roArgThrAlaAspGlySerValAspValHisAlaThrThrGlnMetLeu 286

    820 AACCGCAATGCCGAATATTGGATACGCCGTTTTCCGACGCAGTATCTGTT 869
           |||      :::|||   |||:::|||:::   |||   |||::::::::
    287 AsnAspValValGluArgTrpValArgGluAspProGlyGlnTrpMetTr 303

    870 TATGTACAACCGCTACAAAATG 891
        ::::::::::|||:::::::::
    303 pPheHisLysArgTrpGluIle 310

seq_name: sp_bacteria:O82937

seq_documentation_block:
ID   O82937      PRELIMINARY;      PRT;   332 AA.
AC   O82937;
DT   01-NOV-1998 (TrEMBLrel. 08, Created)
DT   01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   LIPID A BIOSYNTHESIS (KDO)2-(LAUROYL)-LIPID IVA ACYLTRANSFERASE.
GN   ECF4 OR MSBB.
OS   Escherichia coli O157:H7.
OG   Plasmid pO157.
OC   Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC   Escherichia.
OX   NCBI_TaxID=83334;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=98290540; PubMed=9628576;
RA   Makino K., Ishii K., Yasunaga T., Hattori M., Yokoyama K.,
RA   Yatsudo H.C., Kubota Y., Yamaichi Y., Iida T., Yamamoto K., Honda T.,
RA   Han C., Ohtsubo A., Kasamatsu M., Hayashi T., Kuhara S., Shinagawa H.;
RT   "Complete nucleotide sequences of 93-kb and 3.3-kb plasmids of an
RT   enterohemorrhagic Escherichia coli O157:H7 derived from Sakai
RT   outbreak.";
RL   DNA Res. 5:1-9(1998).
RN   [2]
RP   SEQUENCE OF 1-235 FROM N.A.
RC   STRAIN=4304-PT14;
RX   MEDLINE=98261495; PubMed=9596716;
RA   Boerlin P., Chen S., Colbourne J.K., Johnson R., De Grandis S.,
RA   Gyles C.;
RT   "Evolution of enterohemorrhagic Escherichia coli hemolysin plasmids
RT   and the locus for enterocyte effacement in shiga toxin-producing E.
RT   coli.";
RL   Infect. Immun. 66:2553-2561(1998).
RN   [3]
RP   SEQUENCE FROM N.A.
RC   STRAIN=EDL933;
RA   Brunder W.;
RL   Submitted (FEB-1997) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AB011549; BAA31840.1; -.
DR   EMBL; AF043470; AAC24348.1; -.
DR   EMBL; Y11275; CAA72141.1; -.
KW   Transferase; Acyltransferase; Plasmid.
SQ   SEQUENCE   332 AA;  37858 MW;  0D49F6B93E29E9A6 CRC64;


alignment_scores:
            Quality:  151.00              Length:    286
              Ratio:  0.993                Gaps:      9
 Percent Similarity:  53.147   Percent Identity: 22.028

alignment_block:
US-09-303-518D-569 x O82937   ..

Align seg 1/1  to: O82937  from: 1  to: 332

     82 TCCCTGCTGCCGCTTTCCTGTCTGCACACGCTGGGAAACCGGCTCGGACA 131
        :::::::::|||   :::            |||   ::::::||||||
     56 AlaMetValProProAlaLeuArgAspProLeuLeuGlyLysLeuGlyMe 72
```

```
 132 TCTGGCGTTTTACCTTTTAAAGGAAGACCGCGCGCGCATCGTCGCCAATA 181
      |||:::      |||   |||:::   |||   |||   :::   |||:
  72 tLeuValGlyArgLeuGlyLysSerAlaArgGlnArgAlaLeuIleAsnL 89

 182 TGCGT.....CAGGCAGGCATGAATCCCGACCCCAAAACGGTC 219
     ::          :::   :::   :::         :::   |||
  89 euSerLeuCysPheProGluTyrSerAspLysGluLysGluAsnIleVal 105

 220 AAAGCCGTTTTTGCGGAAACGGCAAAAGGCGGTTTGGAACTTGCCCCCGC 269
       |||:::||||||   ::::::   :::   :::   :::|||
 106 AspAlaMetPheAlaThrAlaSerMetAlaValValLeuMetAlaGluLe 122

 270 GTTTTTCAGAAAACCGGAAGACATAGAAACAATGTTCAAAGCGGTACACG 319
       :::      |||:::   |||:::      ::::::      :::|
 122 uAlaLeuSerGlyProAspLysIleSerHisArgIleArg...TrpAsnG 138

 320 GCTGGGAACATGTGCAGCAGGCTTTGGACAAACACGAAGGGCTGCTATTC 369
     ||   |||   |||::::::      :::::::::|||   ::::::|||
 138 lyLeuGluIleValGluLysMetAlaGlnAsnAsnGluLysValIlePhe 154

 370 ATCACGCCGCACATCGGCAGCTACGATTTG..................... 399
     ::::::||||||      :::   |||:::
 155 LeuValProHisAlaTrpGlyValAspIleProAlaMetLeuMetAlaAl 171

 400 .GGCGGACGCTACATCAGCCAGCAGCTTCCGTTCCCGCTGACCGCCATGT 448
      :::||||||                          ::::::||||||:
 171 aSerGlyArg.......................LysMetAlaAlaMetP 180

 449 ACAAACCGCCGAAAATCAAAGCGATAGACAAAATCATGCAGGCGGGCAGG 498
     ::         :::      :::::::|||   :::   ::::::   |||
 180 heHisAsnGlnArgAsnProValValAspTyrValTrpAsnSerValArg 196

 499 GTTCGCGGCAAAGGAAAAACCGCGCCTACCAGCATACAAGGGGTCAAACA 548
       |||      ||||||            :::   :::|||:::   ::
 197 ArgArgPheGlyGlyLysLeuHisAlaArgAsn...AspGlyIleAlaSe 212

 549 AATCATCAAAGCCCTGCGTTCGGGCGAAGCAACCATCGTCCTGCCC.... 594
     ::::::::::::::::|||:::|||            ||||||
 212 rPheValArgSerValArgGlnGlyTyrTrpGlyTyrTyrLeuProAspG 229

 595 ..GACCACGTCCCCTCCCCTCAAGAAGGCGGGGAAGGCGTATGGGTGGAT 642
       ||||||   |||:::   :::|||                :::::|||
 229 lnAspHisGlyProGluPheSerGlu.....:.....PheAlaAsp 240

 643 TTCTTCGGCAAACCTGCCTATACCATGACGCTGGCGGCAAAATTGGCACA 692
     ||||||:::                |||:::   :::   ::::::|||:::::
 241 PhePheAlaThrTyrLysAlaThrLeuProValIleGlyArgLeuSerAr 257

 693 CGTCAAAGGCGTGAAAACCCTGTTTTTCTGCTGCGAACGCCTGCCTGGCG 742
     :::::::|||::::::   :::                    :::|||
 257 gIleSerGlyAlaArgIleIle.....ProLeuPheProValT 270

 743 GACAAGGT.....TTCGATTTGCACATCCGCCCCGTCCAAGGG 780
       :::|||          :::   :::|||:::   |||      :::
 270 yrAspGlyLysThrHisHisLeuThrIleHisValSerProProLeuAla 286

 781 GAATTGAACGGCGACAAAGCCCATGATGCCGCCGTGTTCAACCGCAATGC 830
           :::   :::   ||||||   |||      :::|||:::   ::
 287 IleArgGlnLysSerAspAlaHisIleAlaArgGlnIleAsnGluValVa 303

 831 CGAATATTGGATACGCCGTTTTCCGACGCAGTATCTGTTTATGTACAACC 880
     :|||   ::::::|||      |||   ||||||   ::::::   :::
 303 lGluAsnPheValArgProHisProGluGlnTyrThrTrpIleLeuLysL 320

 881 GCTACAAA 888
           |||
 320 euLeuLys 322
```

```
seq_name: sp_bacteria:Q9ZGU3

seq_documentation_block:
ID   Q9ZGU3      PRELIMINARY;      PRT;   343 AA.
AC   Q9ZGU3;
DT   01-MAY-1999 (TrEMBLrel. 10, Created)
DT   01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT   01-MAY-1999 (TrEMBLrel. 10, Last annotation update)
DE   PUTATIVE ACYLTRANSFERASE.
GN   L7029.
OS   Escherichia coli O157:H7.
OG   Plasmid pO157.
OC   Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC   Escherichia.
OX   NCBI_TaxID=83334;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=EDL933;
RX   MEDLINE=98391744; PubMed=9722640;
RA   Burland V., Shao Y., Perna N.T., Plunkett G., Sofia H.J.,
RA   Blattner F.R.;
RT   "The complete DNA sequence and analysis of the large virulence plasmid
RT   of Escherichia coli O157:H7.";
RL   Nucleic Acids Res. 26:4196-4204(1998).
DR   EMBL; AF074613; AAC70097.1; -.
KW   Transferase; Acyltransferase; Plasmid.
SQ   SEQUENCE   343 AA;  39074 MW;  EB869F0D3D30DB6A CRC64;


alignment_scores:
            Quality:  151.00              Length:    286
              Ratio:  0.993                Gaps:      9
 Percent Similarity:  53.147   Percent Identity: 22.028

alignment_block:
 US-09-303-518D-569 x Q9ZGU3    ..

 Align seg 1/1  to: Q9ZGU3  from: 1  to: 343

  82 TCCCTGCTGCCGCTTTCCTGTCTGCACACGCTGGGAAACCGGCTCGGACA 131
     ::::::::::|||   :::             |||   ::::::||||||
  67 AlaMetValProProAlaLeuArgAspProLeuGluGlyLysLeuGlyMe 83

 132 TCTGGCGTTTTACCTTTTAAAGGAAGACCGCGCGCGCATCGTCGCCAATA 181
      |||:::      |||   |||:::   |||   |||   :::   |||:
  83 tLeuValGlyArgLeuGlyLysSerAlaArgGlnArgAlaLeuIleAsnL 100

 182 TGCGT.....CAGGCAGGCATGAATCCCGACCCCAAAACGGTC 219
     ::          :::   :::   :::         :::   |||
 100 euSerLeuCysPheProGluTyrSerAspLysGluLysGluAsnIleVal 116

 220 AAAGCCGTTTTTGCGGAAACGGCAAAAGGCGGTTTGGAACTTGCCCCCGC 269
       |||:::||||||   ::::::   :::   :::   :::|||
 117 AspAlaMetPheAlaThrAlaSerMetAlaValValLeuMetAlaGluLe 133

 270 GTTTTTCAGAAAACCGGAAGACATAGAAACAATGTTCAAAGCGGTACACG 319
       :::      |||:::   |||:::      ::::::      :::|
 133 uAlaLeuSerGlyProAspLysIleSerHisArgIleArg...TrpAsnG 149

 320 GCTGGGAACATGTGCAGCAGGCTTTGGACAAACACGAAGGGCTGCTATTC 369
     ||   |||   |||::::::      :::::::::|||   ::::::|||
 149 lyLeuGluIleValGluLysMetAlaGlnAsnAsnGluLysValIlePhe 165

 370 ATCACGCCGCACATCGGCAGCTACGATTTG..................... 399
     ::::::||||||      :::   |||:::
 166 LeuValProHisAlaTrpGlyValAspIleProAlaMetLeuMetAlaAl 182

 400 .GGCGGACGCTACATCAGCCAGCAGCTTCCGTTCCCGCTGACCGCCATGT 448
      :::||||||                          ::::::||||||:
 182 aSerGlyArg.......................LysMetAlaAlaMetP 191

 449 ACAAACCGCCGAAAATCAAAGCGATAGACAAAATCATGCAGGCGGGCAGG 498
     ::         :::      :::::::|||   :::   ::::::   |||
 191 heHisAsnGlnArgAsnProValValAspTyrValTrpAsnSerValArg 207

 499 GTTCGCGGCAAAGGAAAAACCGCGCCTACCAGCATACAAGGGGTCAAACA 548
       |||      ||||||            :::   :::|||:::   ::
 208 ArgArgPheGlyGlyLysLeuHisAlaArgAsn...AspGlyIleAlaSe 223
```

```
549 AATCATCAAAGCCCTGCGTTCGGGCGAAGCAACCATCGTCCTGCCC.... 594
    :::::::::::::::::|||:::|||                 ||||||
223 rPheValArgSerValArgGlnGlyTyrTrpGlyTyrTyrLeuProAspG 240

595 ..GACCACGTCCCCTCCCCTCAAGAAGGCGGGGAAGGCGTATGGGTGGAT 642
      ||||||   |||:::   :::|||              :::::::|||
240 lnAspHisGlyProGluPheSerGlu..............PheAlaAsp 251

643 TTCTTCGGCAAACCTGCCTATACCATGACGCTGGCGGCAAAATTGGCACA 692
    ||||||:::                |||:::   :::   ::::::|||:::::
252 PhePheAlaThrTyrLysAlaThrLeuProValIleGlyArgLeuSerAr 268

693 CGTCAAAGGCGTGAAAACCCTGTTTTTCTGCTGCGAACGCCTGCCTGGCG 742
    :::::::|||::::::   :::                  :::|||
268 gIleSerGlyAlaArgIleIle............ProLeuPheProValT 281

743 GACAAGGT............TTCGATTTGCACATCCGCCCCGTCCAAGGG 780
      :::|||           :::   :::|||:::   |||       :::
281 yrAspGlyLysThrHisHisLeuThrIleHisValSerProProLeuAla 297

781 GAATTGAACGGCGACAAAGCCCATGATGCCGCCGTGTTCAACCGCAATGC 830
          :::   :::   ||||||   |||      :::|||:::   ::
298 IleArgGlnLysSerAspAlaHisIleAlaArgGlnIleAsnGluValVa 314

831 CGAATATTGGATACGCCGTTTTCCGACGCAGTATCTGTTTATGTACAACC 880
    :|||   ::::::|||      |||   ||||||   ::::::   :::
314 lGluAsnPheValArgProHisProGluGlnTyrThrTrpIleLeuLysL 331

881 GCTACAAA 888
      |||
331 euLeuLys 333
```

seq_name: sp_bacteriap:Q92GN2

```
seq_documentation_block:
ID   Q92GN2      PRELIMINARY;      PRT;   290 AA.
AC   Q92GN2;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   LIPID A BIOSYNTHESIS LAUROYL ACYLTRANSFERASE.
GN   HTRB OR RC1091.
OS   Rickettsia conorii.
OC   Bacteria; Proteobacteria; alpha subdivision; Rickettsiales;
OC   Rickettsiaceae; Rickettsieae; Rickettsia.
OX   NCBI_TaxID=781;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=MALISH 7;
RX   MEDLINE=21442074; PubMed=11557893;
RA   Ogata H., Audic S., Renesto-Audiffren P., Fournier P.-E., Barbe V.,
RA   Samson D., Roux V., Cossart P., Weissenbach J., Claverie J.-M.,
RA   Raoult D.;
RT   "Mechanisms of evolution in Rickettsia conorii and R. prowazekii.";
RL   Science 293:2093-2098(2001).
DR   EMBL; AE008660; AAL03629.1; -.
KW   Transferase; Acyltransferase; Complete proteome.
SQ   SEQUENCE   290 AA;  33522 MW;  1C0FD05318332FD0 CRC64;
```

```
alignment_scores:
            Quality:  149.00              Length:    201
              Ratio:  1.221                 Gaps:      8
 Percent Similarity:  60.697   Percent Identity: 23.383

alignment_block:
 US-09-303-518D-569 x Q92GN2    ..

 Align seg 1/1  to: Q92GN2  from: 1  to: 290

     310 GCGGTACACGGCTGGGAACATGTGCAGCAGGCTTTGGACAAACACGAAGG 359
         |||:::   |||   |||::::::::::            |||   |||||
```

```
101 AlaIleIleGlyIleGluAsnIleLys...........LysLeuGluGl 113

360 G......CTGCTATTCATCACGCCGCACATCGGCAGCTACGATTTGGGCG 403
    |      :::|||::::::      |||:::::::::::|||::::::
113 yGlnProPheLeuLeuPheGlyGlyHisPheAlaAsnTrpAspIleSerI 130

404 GACGCTACATCAGCCAGCAGCTTCCGTTCCCGCTGACCGCCATGTACAAA 453
      :::   :::   ::       |||      :::::::::::|||:::
130 leLysIleLeuHisLysPheTyrPro...LysValAlaValIleTyrArg 145

454 CCGCCGAAAATCAAAGCGATAGACAAAATCATGCAGGCGGGCAGGGTTCG 503
        :::         ::::::|||:::::::::   :::|||:::
146 LysAlaAsnAsnProTyrValAsnLysLeuValAsnGluSerArgAlaGl 162

504 CGGCAAAGGAAAAACCGCGCCTACCAGCATACAAGGGGTCAAACAAATCA 553
      |||   :::      |||   :::   :::|||   :::   ::::
162 yAspLysLeuArgLeuIleProLysGlyProGluGlySerArgAlaLeuV 179

554 TCAAAGCCCTGCGTTCGGGCGAAGCAACCATCGTCCTGCCCGACCACGTC 603
    :::::|||:::::::::::::|||:::   ::::::|||   |||
179 alArgAlaIleLysGluSerGluSerIleValMetLeuValAsp..... 193

604 CCCTCCCCTCAAGAAGGCGGGGAAGGCGTATGGGTGGATTTCTTCGGCAA 653
          |||:::   ::::::|||:::   |||   |||:::|||
194 .........GlnLysMetAsnAspGlyIleGluValProPheLeuGlyHi 207

654 ACCTGCCTATACCATGACGCTGGCGGCAAAATTGGCACACGTCAAAGGCG 703
      ||||||   |||   :::      |||||||:::|||
207 sProAlaMetThrAlaSerAlaIleAlaLysIleAla......LeuGlnT 222

704 TGAAAACCCTGTTTTTCTGCTGCGAACGCCTG...CCTGGCGGACAAGGT 750
      |||      ::::::   |||:::   :::         |||:::
222 yrLysTyrProIleIleProCysGlnIleIleArgThrLysGlySerTyr 238

751 TTCGATTTGCACATCCGCCCC...GTCCAAGGGGAATTGAACGGCGACAA 797
    |||   :::   ::::::|||   ::::::   |||   :::|||||::
239 PheLysValIleValHisProGlnLeuLysPheGluGlnThrGlyAspAs 255

798 AGCCCATGATGCC.............GCCGTGTTCAACCGCAATGCCGAAT 835
    :      |||:::                    :::|||:::
255 nLysAlaAspCysTyrAsnIleMetLeuAsnIleAsnGlnMetLeuGlyG 272

836 ATTGGATACGCCGTTTTCCGACGCAGTATCTGTTTATGTACAACCGCTAC 885
      |||:::::::::   |||:::|||::::::::::::::::||||||:::
272 luTrpValLysGlnAsnProAlaGlnTrpPheTrpPheHisAsnArgTrp 288

886 AAA 888
    |||
289 Lys 289
```

seq_name: sp_bacteria:P94909

```
seq_documentation_block:
ID   P94909      PRELIMINARY;      PRT;   529 AA.
AC   P94909;
DT   01-MAY-1997 (TrEMBLrel. 03, Created)
DT   01-MAY-1997 (TrEMBLrel. 03, Last sequence update)
DT   01-NOV-1998 (TrEMBLrel. 08, Last annotation update)
DE   HYPOTHETICAL 58.7 KDA PROTEIN (FRAGMENT).
OS   Microbacterium ammoniaphilum.
OC   Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC   Actinomycetales; Micrococcineae; Microbacteriaceae; Microbacterium.
OX   NCBI_TaxID=42460;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=ATCC 15354;
RX   MEDLINE=96257250; PubMed=8654988;
RA   Striebel H.M., Seeber S., Jarsch M., Kessler C.;
RT   "Cloning and characterization of the MamI restriction-modification
RT   system from Microbacterium ammoniaphilum in Escherichia coli.";
RL   Gene 172:41-46(1996).
DR   EMBL; X79027; CAA55649.1; -.
KW   Hypothetical protein.
```

```
FT   NON_TER       1      1
SQ   SEQUENCE   529 AA;  58716 MW;  5DC1216237DF772D CRC64;


alignment_scores:
            Quality:  147.50              Length:    293
              Ratio:  1.209                Gaps:     16
 Percent Similarity:  41.638   Percent Identity: 28.669

alignment_block:
 US-09-303-518D-569 x P94909   ..

 Align seg 1/1  to: P94909  from: 1  to: 529

     30 CCCTTTGCGAACCGCCATGCACATCCTGTTGACCGCCCTGCTCAAATGCC 79
        |||      :::|||::::::|||||::::::|||||   |||   ||
     74 ProAlaGlnArgArgGlnValHisProLeuGlnArgProHisGlnGluPr 90

     80 T.................................................. 80
        |
     90 oGlyAlaArgGlyGluLeuProValArgAspAspArgAlaGluArgGlyA 107

     81 .............CTCCCTGCTGCCGCTTTCCTGTCTGCACAC...... 110
                     |||||||||          |||
    107 rgArgGluProAlaArgProAlaAlaArgAspAlaArgArgHisLeuArg 123

    111 ...GCTGGGAAACCGGCTCGGACATCTGGCGTTTTACCTTTTAAAGGAAG 157
           ::::::   |||||||||      |||||||||
    124 LeuGlyAlaHisProAlaArg...ArgGlyValLeu.............. 134

    158 ACCGCGCGCGCATCGTCGCCAATATGCGTCAGGCAGGCATGAATCCCGAC 207
           |||   ||||||||||:::   |||   |||::::::          |
    135 ....ArgArgHisArgArgHisArgAlaArgGlyGluArgGlyArgGlyP 150

    208 CCCAAAACGGTCAAAGCCGTTTTTGCGGAAACGGCAAAAGG...CGGTTT 254
        ||::::::         |||         :::|||:::|||   |||
    150 roArgGlnGlnValProArgGlnHisProArgGlyArgArgAspArgAla 166

    255 GGAACTTGCCCCCGCGTTTTTCAGAAAACCGGAAGACATAGAAACAATGT 304
        |||                    :::   |||   ||||||::::::
    167 Gly.....................ArgProGlyLeuHisArgArgArgAr 176

    305 TCAAAGCGGTACACGGCTGGGAACATGTGCAGCAGGCTTTGGACAAACAC 354
         ::::::      |||   |||         :::      |||:::   :
    176 gArgAlaArgArgArgGlnGlyArgProGlnValArgHisGlyAspAspG 193

    355 GAAGGGCTGCTATTCATCACGCCGCACATCGGCAGCTACGATTTGGGCGG 404
        ::          |||:::|||      |||:::   |||      |||
    193 ln............HisArgAlaAspProArgArgProArgAspProArg 205

    405 ACGCTACATCAGCCAGCAGCTTCCGTTCCCGCTGACCGCCATGTACAAAC 454
        :::   |||:::|||
    206 AlaGlyHisHisProLeu.............................. 211

    455 CGCCGAAAATCAAAGCGATAGACAAAATCATGCAGGCGGGCAGGGTTCGC 504
                                  :::   ||||||||||||::::
    212 ............................ArgGluGlyGlyGlnGlyGlnG 219

    505 GGCAAAGGAAAAACCGCGCCTAC......CAGCATACAAGGGGTCAAACA 548
        ::   |||      ||||||         |||         |||
    219 luAlaArgProLeuArgAlaArgGlyGlyGlnGlyGlyProGlyGlyPro 235

    549 AATCATCAAAGCCCTGCGTTCGGG........................... 572
              :::   ||||||:::|||
    236 AlaGlyArgHisProAlaLeuGlyAspGluAspArgProArgAlaArgGl 252

    573 .......CGAAGCAACCATCGTCCTGCCCGACCACGTCCCCTCCCCTCAA 615
              |||:::   ||||||  |||                |||   :::|
    252 nGlyAlaArgAlaAlaHisArgGluAlaLeuHisLeuArgLeu...GlnA 268

    616 GAAGGCGGGGAAGG............CGTATGGGTGGATTTCTTCGGCAA 653
        ||||||||||||||            |||      |||      |||:::
    268 rgArgArgGlyArgAlaAspGlyCysArgAlaGlnGlyArgAlaArgArg 284

    654 ACCTGCCTATACCATGACGCTGGCGGCAAAATTGGCACACGTCAAAGGCG 703
        :::         :::   |||||||||      |||         |||||
    285 Ala.........ArgArgAlaGlyGly......GlyGlyLeuProArgAr 296

    704 TGAAAACCCTGTTTTTCTGCTGCGAACGCCTGCCTGGCGGACAAGGTTTC 753
        ||||:::                              |||:::|||
    296 gGluAsp.............................ArgValArg.... 301

    754 GATTTGCACATCCGCCCCGTCCAAGGGGA 782
              ||||||   ||||||||||||||||
    302 .....AlaHisArgProArgProArgGly 309

seq_name: sp_bacteria:Q9AKK5

seq_documentation_block:
ID   Q9AKK5       PRELIMINARY;      PRT;   290 AA.
AC   Q9AKK5;
DT   01-JUN-2001 (TrEMBLrel. 17, Created)
DT   01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   HTRB PROETIN.
GN   HTRB.
OS   Rickettsia montana.
OC   Bacteria; Proteobacteria; alpha subdivision; Rickettsiales;
OC   Rickettsiaceae; Rickettsieae; Rickettsia.
OX   NCBI_TaxID=33991;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=21219194; PubMed=11319266;
RA   Andersson J.O., Andersson S.G.E.;
RT   "Pseudogenes, junk DNA and the dynamics of Rickettsia genomes.";
RL   Mol. Biol. Evol. 18:829-839(2001).
DR   EMBL; AJ293330; CAC33650.1; -.
SQ   SEQUENCE   290 AA;  33635 MW;  9599E3C0C3C076F9 CRC64;


alignment_scores:
            Quality:  147.00              Length:    190
              Ratio:  1.267                Gaps:      7
 Percent Similarity:  61.053   Percent Identity: 23.158

alignment_block:
 US-09-303-518D-569 x Q9AKK5   ..

 Align seg 1/1  to: Q9AKK5  from: 1  to: 290

    343 TTGGACAAACACGAAGGG......CTGCTATTCATCACGCCGCACATCGG 386
        :::   |||   ||||||      :::|||:::::::::   |||:::::
    108 IleLysLysLeuGluGlyGlnProPheLeuLeuPheSerGlyHisPheAl 124

    387 CAGCTACGATTTGGGCGGACGCTACATCAGCCAGCAGCTTCCGTTCCCGC 436
        :::::::|||::::::   :::   :::   :::      |||      :
    124 aAsnTrpAspIleSerLeuLysLeuLeuHisLysPheTyrPro...LysV 140

    437 TGACCGCCATGTACAAACCGCCGAAAATCAAAGCGATAGACAAAATCATG 486
        ::::::::::|||:::      :::         ::::::|||::::::
    140 alAlaValIleTyrArgLysAlaAsnAsnProTyrValAsnLysLeuVal 156

    487 CAGGCGGGCAGGGTTCGCGGCAAAGGAAAAACCGCGCCTACCAGCATACA 536
        :::   :::|||:::      |||   :::      |||   :::   ::
    157 AsnGluSerArgAlaGlyAspLysLeuArgLeuIleProLysGlyProGl 173

    537 AGGGGTCAAACAAATCATCAAAGCCCTGCGTTCGGGCGAAGCAACCATCG 586
        :|||   :::   :::::::::|||::::::::::::|||      ::::
    173 uGlySerArgAlaLeuValArgAlaIleLysGluSerGluProIleValM 190

    587 TCCTGCCCGACCACGTCCCCTCCCCTCAAGAAGGCGGGGAAGGCGTATGG 636
        ::|||   |||              |||:::   ::::::|||:::
    190 etLeuValAsp.............GlnLysMetAsnAspGlyIleGlu 201

    637 GTGGATTTCTTCGGCAAACCTGCCTATACCATGACGCTGGCGGCAAAATT 686
        |||  .|||:::|||   ||||||   |||   :::      ||||||::
```

```
      202 ValProPheLeuGlyHisProAlaMetThrAlaSerAlaIleAlaLysIl 218

      687 GGCACACGTCAAAGGCGTGAAAACCCTGTTTTTCTGCTGCGAACGCCTG. 735
          :|||               |||      ::::::   |||:::   :::
      218 eAla......LeuGlnTyrLysTyrProIleIleProCysGlnIleIleA 233

      736 ..CCTGGCGGACAAGGTTTCGATTTGCACATCCGCCCC...GTCCAAGGG 780
                ||| :::   |||   :::   :::::::|||   ::::::
      233 rgThrLysGlySerTyrPheLysValIleValHisProGlnLeuLysPhe 249

      781 GAATTGAACGGCGACAAAGCCCATGATGCC.............GCCGTGTT 818
          |||   :::||||||:::       |||:::                ::
      250 GluGlnThrGlyAspAsnLysAlaAspCysTyrAsnIleMetLeuAsnIl 266

      819 CAACCGCAATGCCGAATATTGGATACGCCGTTTTCCGACGCAGTATCTGT 868
          :|||:::            |||:::::::::   |||:::|||:::::::
      266 eAsnGlnMetLeuGlyGluTrpValLysGlnAsnProSerGlnTrpPheT 283

      869 TTATGTACAACCGCTACAAA 888
          ::::::::||||||:::|||
      283 rpPheHisAsnArgTrpLys 289

seq_name: sp_bacteriap:Q9K217

seq_documentation_block:
ID   Q9K217      PRELIMINARY;      PRT;   462 AA.
AC   Q9K217;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-OCT-2001 (TrEMBLrel. 18, Last annotation update)
DE   LIPID A BIOSYNTHESIS LAUROYL ACYLTRANSFERASE, PUTATIVE.
GN   CP0676.
OS   Chlamydia pneumoniae (Chlamydophila pneumoniae).
OC   Bacteria; Chlamydiales; Chlamydiaceae; Chlamydophila.
OX   NCBI_TaxID=83558;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=AR39;
RX   MEDLINE=20150255; PubMed=10684935;
RA   Read T.D., Brunham R.C., Shen C., Gill S.R., Heidelberg J.F.,
RA   White O., Hickey E.K., Peterson J., Utterback T., Berry K., Bass S.,
RA   Linher K., Weidman J., Khouri H., Craven B., Bowman C., Dodson R.,
RA   Gwinn M., Nelson W., DeBoy R., Kolonay J., McClarty G., Salzberg S.L.,
RA   Eisen J., Fraser C.M.;
RT   "Genome sequences of Chlamydia trachomatis MoPn and Chlamydia
RT   pneumoniae AR39.";
RL   Nucleic Acids Res. 28:1397-1406(2000).
DR   EMBL; AE002225; AAF38487.1; -.
DR   TIGR; CP0676; -.
DR   InterPro; IPR000504; RRM.
DR   PROSITE; PS00030; RRM_RNP_1; UNKNOWN_1.
KW   Transferase; Acyltransferase.
SQ   SEQUENCE   462 AA;  52624 MW;  C31CD0C9B01E13DE CRC64;


alignment_scores:
            Quality:  146.00              Length:    331
              Ratio:  0.869                Gaps:     13
 Percent Similarity:  50.755   Percent Identity: 22.356

alignment_block:
 US-09-303-518D-569 x Q9K217   ..

 Align seg 1/1  to: Q9K217  from: 1  to: 462

      34 TTGCGAACCGCCATGCACATCCTGTTGACCGCCCTGCTCAAATGCCTCTC 83
         |||::::::   ::::::   |||:::::::::::::::
      13 LeuGluAlaProLeuTyrTyrLeuValSerGlyIleIleAlaLeuCysAr 29

      84 CCTGCTGCCGCTTTCCTGTCTGCACACGCTGGGAAACCGGCTCGGACATC 133
            |||   |||   |||      ||||||:::   :::|||   |
      29 gHisThrProArgSerPheLeuThrGlyLeuGlyLysGlyPheGlyPheL 46

     134 TGGCGTTTTACCTTTTAAAGGAAGACCGCGCGCGCATCGTCGCCAATATG 183
         ||||||||||:::::::::::::   |||         ::::::|||:::
      46 euAlaPheTyrIleIleSerAspTyrArgLysThrAlaLeuThrAsnLeu 62

     184 ..............................................CG 185
                                                      ||
      63 AlaLeuAlaPheProGluLysThrPheAspGluArgTyrLysIleAlaAr 79

     186 TCAGGCAGGCATGAATCCCGACCCCAAAACGGTCAAAGCCGTTTTTGCGG 235
         ||||:::   ::::::                ::::::   :::      |
      79 gGlnSerLeuGlnHisLeuIleIleThrLeuLeuGluLeuLeuAlaIleG 96

     236 AAACGGCAAAAGGCGGTTTGGAACTTGCCCCCGCGTTTTTC..... 276
         ||        |||::::::::::        ::::::
      96 luGlnLeuValGlyAsnIleAspLysLeuIleThrIleValThrSerSer 112

     277 AGAAAACCG.............................GAAGACAT 293
         |||:::|||                             |||||::
     113 ArgAsnProLysGlyPheSerSerGluGluValIleSerAsnGluAspLe 129

     294 AGAAACAATGTTCAAAGCGGTACACGGCTGGGAACATGTGCAGCAGGCTT 343
         :|||      ||||||                              |
     129 uGluGluThrPheLys.............................AsnL 136

     344 TGGACAAACACGAAGGGCTGCTATTCATCACGCCGCACATCGGCAGCTAC 393
         ||::::::   :::||||||::::::::        |||   :::::::::
     136 euGlnGluLysGlnGlyLeuIleLeuPheCysGlyHisGlnAlaAsnTrp 152

     394 GATTTGGGCGGACGCTACATCAGCCAGCAGCTTCCGTTCCCGCTGACCGC 443
         :::|||        ||||||:::::::::       :::|||       ||
     153 GluLeuProPheLeuTyrIleThrLysAsn......TyrProGlyIleAl 167

     444 CATGTACAAACCGCCGAAAATCAAAGCGATAGACAAAATCATGCAGGCGG 493
         |:::   |||      |||   :::   :::::::|||   :::   |||
     167 aPheAlaLysAlaIleLysAsnGlnArgLeuSerLysLysIlePheAlaL 184

     494 GCAGGGTTCGCGGCAAAGGAAAAACCGCGCCTACCAGCATACAAGGGGTC 543
           |||         |||||||||   :::|||   :::   :::|||:::
     184 euArgGluValPheLysGlyLysIleValProProLys...AsnGlyIle 199

     544 AAACAAATCATCAAAGCCCTGCGTTCGGGCGAAGCAACCATCGTCCTGCC 593
         :::|||   |||:::||||||::::::|||:::   :::   ::::::
     200 GlnGlnGlyIleGluAlaLeuAsnGlnGlyLysLeuValGlyIleValGl 216

     594 CGACCACGTC.....................CCCTCCCCTCAAGAAG 619
          |||::::::                     |||
     216 yAspGlnAlaLeuLeuMetSerSerTyrThrTyrPro............. 228

     620 GCGGGGAAGGCGTATGGGTGGATTTCTTCGGCAAACCTGCCTATACCATG 669
                                 :::||||||:::||||||:::|||
     229 .....................LeuPheGlySerProAlaPheThrThr 237

     670 ACGCTGGCGGCAAAATTGGCACACGTCAAAGGCGTGAAAACCCTGTTTTT 719
         |||      |||   ||||||:::      |||      ::::::
     238 ThrSerProAlaLeuLeuAlaTyrLysThrGlyProValIleAlaVa 254

     720 CTGCTGCGAACGCCTGCCTGGCGGACAAGGTTTCGATTTGCACATCCGCC 769
               :::|||         ::::::||||||:::      :::   |
     254 lAsnValSerArg......GlnAlaLysGlyPheGlu......ValIleP 267

     770 CCGTCCAAGGGGAATTGAACGGCGACAAAGCC............CATGAT 807
         ||   ::::::::::|||   ::::::|||:::                ::::::
     267 ro...SerAlaLysLeuTyrAlaAsnLysSerLeuProMetLysGluSer 282

     808 GCCGCCGTGTTCAACCGCAAT............GCCGAATATTGGATACG 845
         :::|||::::::      :::               |||      |||
     283 ValAlaIleLeuMetAspGlnMetMetGlyPheLeuGluLysGlyIleAl 299

     846 CCGTTTTCCGACGCAGTATCTGTTTATGTACAACCGCTACAAA 888
               |||   |||:::::::::::::::::::|||:::|||
     299 aSerGlnProGluGlnTrpMetTrpIleHisLysArgTrpLys 313

seq_name: sp_bacteriap:Q9Z983
```

```
seq_documentation_block:
ID   Q9Z983      PRELIMINARY;      PRT;   467 AA.
AC   Q9Z983;
DT   01-MAY-1999 (TrEMBLrel. 10, Created)
DT   01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   ACYLTRANSFERASE.
GN   HTRB OR CPN0098.
OS   Chlamydia pneumoniae (Chlamydophila pneumoniae).
OC   Bacteria; Chlamydiales; Chlamydiaceae; Chlamydophila.
OX   NCBI_TaxID=83558;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=CWL029;
RX   MEDLINE=99206606; PubMed=10192388;
RA   Kalman S., Mitchell W., Marathe R., Lammel C., Fan J., Hyman R.W.,
RA   Olinger L., Grimwood J., Davis R.W., Stephens R.S.;
RT   "Comparative genomes of Chlamydia pneumoniae and C. trachomatis.";
RL   Nat. Genet. 21:385-389(1999).
DR   EMBL; AE001596; AAD18251.1; -.
KW   Transferase; Complete proteome.
SQ   SEQUENCE   467 AA;  53193 MW;  D3C7C284E9220DD0 CRC64;


alignment_scores:
            Quality:  146.00              Length:    331
              Ratio:  0.869                Gaps:     13
 Percent Similarity:  50.755   Percent Identity: 22.356

alignment_block:
 US-09-303-518D-569 x Q9Z983   ..

 Align seg 1/1  to: Q9Z983  from: 1  to: 467

      34 TTGCGAACCGCCATGCACATCCTGTTGACCGCCCTGCTCAAATGCCTCTC 83
         |||::::::   ::::::   |||:::::::::::::::
      13 LeuGluAlaProLeuTyrTyrLeuValSerGlyIleIleAlaLeuCysAr 29

      84 CCTGCTGCCGCTTTCCTGTCTGCACACGCTGGGAAACCGGCTCGGACATC 133
            |||   |||   |||      ||||||:::   :::|||   |
      29 gHisThrProArgSerPheLeuThrGlyLeuGlyLysGlyPheGlyPheL 46

     134 TGGCGTTTTACCTTTTAAAGGAAGACCGCGCGCGCATCGTCGCCAATATG 183
         ||||||||||||:::::::::::   |||         ::::::|||:::
      46 euAlaPheTyrIleIleSerAspTyrArgLysThrAlaLeuThrAsnLeu 62

     184 ..................................................CG 185
                                                          ||
      63 AlaLeuAlaPheProGluLysThrPheAspGluArgTyrLysIleAlaAr 79

     186 TCAGGCAGGCATGAATCCCGACCCCAAAACGGTCAAAGCCGTTTTTGCGG 235
         ||||:::   ::::::                ::::::   :::.     |
      79 gGlnSerLeuGlnHisLeuIleIleThrLeuLeuGluLeuLeuAlaIleG 96

     236 AAACGGCAAAAGGCGGTTTGGAACTTGCCCCCGCGTTTTTC......... 276
         ||        |||::::::::::         ::::::
      96 luGlnLeuValGlyAsnIleAspLysLeuIleThrIleValThrSerSer 112

     277 AGAAAACCG...............................GAAGACAT 293
         |||:::|||                               ||||||::
     113 ArgAsnProLysGlyPheSerSerGluGluValIleSerAsnGluAspLe 129

     294 AGAAACAATGTTCAAAGCGGTACACGGCTGGGAACATGTGCAGCAGGCTT 343
         :|||      ||||||                                 |
     129 uGluGluThrPheLys...........................AsnL 136

     344 TGGACAAACACGAAGGGCTGCTATTCATCACGCCGCACATCGGCAGCTAC 393
         ||::::::   :::||||||::::::::::      |||   :::::::::
     136 euGlnGluLysGlnGlyLeuIleLeuPheCysGlyHisGlnAlaAsnTrp 152

     394 GATTTGGGCGGACGCTACATCAGCCAGCAGCTTCCGTTCCCGCTGACCGC 443
         :::|||         ||||||::::::::::      :::|||      ||

     153 GluLeuProPheLeuTyrIleThrLysAsn......TyrProGlyIleAl 167

     444 CATGTACAAACCGCCGAAAATCAAAGCGATAGACAAAATCATGCAGGCGG 493
         |:::   |||      |||   :::   ::::::|||   :::   |||
     167 aPheAlaLysAlaIleLysAsnGlnArgLeuSerLysLysIlePheAlaL 184

     494 GCAGGGTTCGCGGCAAAGGAAAAACCGCGCCTACCAGCATACAAGGGGTC 543
           |||         |||||||||   :::|||   :::   :::|||:::
     184 euArgGluValPheLysGlyLysIleValProProLys...AsnGlyIle 199

     544 AAACAAATCATCAAAGCCCTGCGTTCGGGCGAAGCAACCATCGTCCTGCC 593
         :::|||   |||:::||||||::::::|||:::   :::   ::::::
     200 GlnGlnGlyIleGluAlaLeuAsnGlnGlyLysLeuValGlyIleValGl 216

     594 CGACCACGTC.......................CCCTCCCCTCAAGAAG 619
          |||::::::                       |||
     216 yAspGlnAlaLeuLeuMetSerSerTyrThrTyrPro..... 228

     620 GCGGGGAAGGCGTATGGGTGGATTTCTTCGGCAAACCTGCCTATACCATG 669
                               :::||||||:::||||||:::|||
     229 .....................LeuPheGlySerProAlaPheThrThr 237

     670 ACGCTGGCGGCAAAATTGGCACACGTCAAAGGCGTGAAAACCCTGTTTTT 719
         |||      |||   ||||||:::      |||      ::::::
     238 ThrSerProAlaLeuLeuAlaTyrLysThrGlyProValIleAlaVa 254

     720 CTGCTGCGAACGCCTGCCTGGCGGACAAGGTTTCGATTTGCACATCCGCC 769
               :::|||         ::::::||||||::      :::   |
     254 lAsnValSerArg......GlnAlaLysGlyPheGlu......ValIleP 267

     770 CCGTCCAAGGGGAATTGAACGGCGACAAAGCC............CATGAT 807
         ||   :::::::::|||   ::::::|||:::                ::::::
     267 ro...SerAlaLysLeuTyrAlaAsnLysSerLeuProMetLysGluSer 282

     808 GCCGCCGTGTTCAACCGCAAT............GCCGAATATTGGATACG 845
         :::|||::::::      :::                |||      |||
     283 ValAlaIleLeuMetAspGlnMetMetGlyPheLeuGluLysGlyIleAl 299

     846 CCGTTTTCCGACGCAGTATCTGTTTATGTACAACCGCTACAAA 888
               |||   |||:::::::::::::::::::|||:::|||
     299 aSerGlnProGluGlnTrpMetTrpIleHisLysArgTrpLys 313

seq_name: sp_bacteriap:Q9KVD4

seq_documentation_block:
ID   Q9KVD4      PRELIMINARY;      PRT;   325 AA.
AC   Q9KVD4;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   LIPID A BIOSYNTHESIS (KDO)2-(LAUROYL)-LIPID IVA ACYLTRANSFERASE.
GN   VC0212.
OS   Vibrio cholerae.
OC   Bacteria; Proteobacteria; gamma subdivision; Vibrionaceae; Vibrio.
OX   NCBI_TaxID=666;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=EL TOR N16961 / SEROTYPE O1;
RX   MEDLINE=20406833; PubMed=10952301;
RA   Heidelberg J.F., Eisen J.A., Nelson W.C., Clayton R.A., Gwinn M.L.,
RA   Dodson R.J., Haft D.H., Hickey E.K., Peterson J.D., Umayam L.A.,
RA   Gill S.R., Nelson K.E., Read T.D., Tettelin H., Richardson D.,
RA   Ermolaeva M.D., Vamathevan J., Bass S., Qin H., Dragoi I., Sellers P.,
RA   McDonald L., Utterback T., Fleischmann R.D., Nierman W.C., White O.,
RA   Salzberg S.L., Smith H.O., Colwell R.R., Mekalanos J.J., Venter J.C.,
RA   Fraser C.M.;
RT   "DNA sequence of both chromosomes of the cholera pathogen Vibrio
RT   cholerae.";
RL   Nature 406:477-483(2000).
DR   EMBL; AE004111; AAF93388.1; -.
DR   TIGR; VC0212; -.
KW   Complete proteome.
SQ   SEQUENCE   325 AA;  37215 MW;  3F949ED94B5C5F92 CRC64;
```

```
alignment_scores:
            Quality:  141.00              Length:    186
              Ratio:  1.270                Gaps:      6
 Percent Similarity:  59.677   Percent Identity: 25.269

alignment_block:
 US-09-303-518D-569 x Q9KVD4   ..

 Align seg 1/1  to: Q9KVD4  from: 1  to: 325

      319 GGCTGGGAACATGTGCAGCAGGCTTTGGACAAACACGAAGGGCTGCTATT 368
          |||   |||::::::          |||||||::   |||   ::::::::
      122 GlyGluGluAsnLeuLeuProLeuLeuAspSerGlyGluArgValIleLe 138

      369 CATCACGCCGCACATCGGCAGCTACGATTTGGGCGGACGCTACATCAGCC 418
          :::::::||||||      :::   |||   ::::::      ::::::
      138 uLeuValProHisSerTrpAlaIleAspTyrAlaAlaValMetLeuAlaA 155

      419 AGCAGCTTCCGTTCCCGCTGACCGCCATGTACAAACCGCCGAAAATCAAA 468
            :::      :::   ::::::   :::   ||||||   :::
      155 laArg...GlyTyrLysValAlaAsnIleMetLysProGlnArgAsnPro 170

      469 GCGATAGACAAAATCATGCAGGCGGGCAGGGTTCGCGGCAAAGGAAAAAC 518
                |||   :::|||::::::   |||::::::      |||:::
      171 IleAlaAspTrpLeuMetHisValGlnArgMetGlnTyrGlyGlyArgIl 187

      519 CGCGCCTACCAGCATACAAGGGGTCAAACAAATCATCAAAGCCCTGCGTT 568
                |||      :::|||:::|||   :::::::::::::::::|
      187 ePhe...ThrArgGluSerGlyIleLysProPheLeuArgSerIleGlnS 203

      569 CGGGCGAAGCAACCATCGTCCTGCCCGACCACGTCCCCTCCCCTCAAGAA 618
          |||||   :::      :::|||||:::      :::||||||
      203 erGlyTyrValGlyTyrTrpValProAspGluAspHisGlyProGln... 218

      619 GGCGGGGAAGGCGTATGGGTGGATTTCTTCGGCAAACCTGCCTATACCAT 668
              ::::::|||:::|||   ||||||:::            |||::
      219 ......AsnSerValPheValProPhePheAlaThrGluLysAlaThrLe 233

      669 GACGCTGGCGGCAAAATTGGCACACGTC...AAAGGCGTGAAAACCCTGT 715
          :         :::|||:::|||   :::   |||:::      :::   :
      233 uLysGlyPheGlyLysMetAlaLysLeuCysLysAlaHisValValProL 250

      716 TTTTCTGCTGCGAACGCCTGCCTGGCGGACAAGGTTTCGATTTGCACATC 765
          :::::   |||   :::      :::|||:::   :::::::::||||||
      250 euMetSerCysTyrAsnSerAspSerGlyArg...TyrGluValHisIle 265

      766 CGCCCC...GTCCAAGGGGAATTGAACGGCGACAAAGCCCATGATGCCGC 812
             |||   :::|||:::      :::|||||::      ||||||
      266 LeuProAlaLeuGlnAsnPheProThrGlyAspGluGluAlaAspAlaLe 282

      813 CGTGTTCAACCGCAATGCCGAATATTGGATACGCCGTTTTCCGACGCAGT 862
          ::::::||||||      |||      :::         |||   ||||
      282 uAlaMetAsnArgAlaIleGluAlaLeuValThrProGlnProGluGlnT 299

      863 ATCTGTTT 870
          ||::::::
      299 yrMetTrp 301

seq_name: sp_bacteriap:Q9KVD3

seq_documentation_block:
ID   Q9KVD3      PRELIMINARY;      PRT;   318 AA.
AC   Q9KVD3;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   LIPID A BIOSYNTHESIS LAUROYL ACYLTRANSFERASE.
GN   VC0213.
OS   Vibrio cholerae.
OC   Bacteria; Proteobacteria; gamma subdivision; Vibrionaceae; Vibrio.
OX   NCBI_TaxID=666;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=EL TOR N16961 / SEROTYPE O1;
RX   MEDLINE=20406833; PubMed=10952301;
RA   Heidelberg J.F., Eisen J.A., Nelson W.C., Clayton R.A., Gwinn M.L.,
RA   Dodson R.J., Haft D.H., Hickey E.K., Peterson J.D., Umayam L.A.,
RA   Gill S.R., Nelson K.E., Read T.D., Tettelin H., Richardson D.,
RA   Ermolaeva M.D., Vamathevan J., Bass S., Qin H., Dragoi I., Sellers P.,
RA   McDonald L., Utterback T., Fleischmann R.D., Nierman W.C., White O.,
RA   Salzberg S.L., Smith H.O., Colwell R.R., Mekalanos J.J., Venter J.C.,
RA   Fraser C.M.;
RT   "DNA sequence of both chromosomes of the cholera pathogen Vibrio
RT   cholerae.";
RL   Nature 406:477-483(2000).
DR   EMBL; AE004111; AAF93389.1; -.
DR   TIGR; VC0213; -.
KW   Transferase; Acyltransferase; Complete proteome.
SQ   SEQUENCE   318 AA;  36542 MW;  FE95D7A4C83106E1 CRC64;


alignment_scores:
            Quality:  136.50              Length:    317
              Ratio:  0.858                Gaps:      15
 Percent Similarity:  50.158   Percent Identity: 24.921

alignment_block:
 US-09-303-518D-569 x Q9KVD3   ..

 Align seg 1/1  to: Q9KVD3  from: 1  to: 318

       58 TTGACCGCCCTGCTCAAATGCCTCTCCCTGCTGCCGCTTTCCTGTCTGCA 107
          |||   |||||||||      ::::::|||||||||          |||
       32 LeuLeuAlaLeuLeu......ValAsnLeuLeuProTyrProValLeuLe 46

      108 CACGCTGGGAAACCGGCTCGGA..........................  129
              :::|||:::   ||||||
       46 uLysIleGlyArgGlyLeuGlyGlnPheSerMetArgPheGlyLysLysA 63

      130 ..........................CATCTGGCGTTTTACCTTTTAAAG 153
                                     :::|||||||||      ::::::
       63 rgValHisIleAlaArgArgAsnLeuGluLeuAlaProPheProThrMetSer 79

      154 GAAGACCGC......GCGCGCATCGTCGCCAATATGCGTCAGGCAGGCAT 197
          :::::::::      |||   ::::::   |||:::::::::::|||
       80 GlnSerGluIleAspAlaPheValLeuGluAsnPheLysAsnThrGly.. 95

      198 GAATCCCGACCCCAAAACGGTCAAAGCCGTTTTTGCGGAAACGGCAAAAG 247
                                 |||::::::   ||||||:::
       96 ........................AlaAlaLeuIleGluThrGlyIleT 104

      248 GCGGTTTGGAACTTGCCCCCGCGTTTTTCAGAAAACCGGAAGACATAGAA 297
              :::      :::      |||   |||          |||   :::
      104 hrTrpPheTrpProThrTrpArgPheLysArgIleLeuIleAspLysAsp 120

      298 ACAATGTTCAAAGCGGTACACGGCTGGGAACATGTGCAGCAGGCTTTGGA 347
          |||                                     |||||||:::
      121 Thr.................................GlnAlaIleAr 125

      348 CAAACAC............GAAGGGCTGCTATTCATCACGCCGCACATCG 385
           :::|||            :::|||:::|||:::         |||
      125 gGlnHisAlaLysThrGlyGlnGlyValLeuLeuCysCysValHisAlaL 142

      386 GCAGCTACGATTTGGGCGGACGCTACATCAGCCAGCAGCTTCCGTTCCCG 435
            :::   ::::::   :::|||   ::::::      |||   :::
      142 euAsnLeuGluIleThrAlaArgAlaPheAla...ValLeuGlyIleGly 157

      436 CTGACCGCCATGTACAAACCGCCGAAAATCAAAGCGATAGACAAAATCAT 485
                ::::::|||:::|||   :::      |||   :::   |||::
      158 GlyTyrGlyValTyrArgProHisSerAsnProAlaTyrGluPheIleGl 174

      486 GCAGGCGGGCAGGGTTCGCGGCAAAGGAAAAACCGCGCCTACCAGCATAC 535
          :      ||||||:::|||:::   ::::::           :::
      174 nTyrArgGlyArgThrArgAsnGlyAsnGlnLeuIleAsnArgThr.... 189
```

```
    536 AAGGGGTCAAACAAATCATCAAAGCCCTGCGTTCGGGCGAAGCAACCATC 585
        :::||||||:::|||:::::::|||||||:::||||||      :::
    190 ..AspIleLysGlnMetIleArgValLeuArgGlnGlyGluArgLeuPhe 205

    586 GTCCTGCCCGACCACGTCCCCTCCCCTCAAGAAGGCGGG......GAAGG 629
           |||||||||              |||:::   |||      :::::
    206 TyrLeuProAsp..............GlnAspTyrGlyHisAsnLysSe 217

    630 CGTATGGGTGGATTTCTTCGGC...AAACCTGCCTATACCATGACGCTGG 676
        :|||:::|||   ||||||:::   :::   |||   |||   |||   :
    217 rValPheValProPhePheAlaValGluGluAlaCysThrThrThrGlyT 234

    677 CGGCAAAATTGGCACACGTCAAAGGCGTGAAAACCCTGTTTTTCTGCTGC 726
        :::::   ||||||:::::::::           ::::::
    234 hrSerIleLeuAlaTyrThrSerHisCysAlaIleValIle......... 247

    727 GAACGCCTGCCTGGCGGACAAGGTTTCGATTTGCACATCCGCCCCGTCCA 776
                     |||:::|||||||            |||   :::||
    248 ..............GlySerGlyPhe............ArgAsnAlaGl 255

    777 AGGG......GAATTGAACGGCGACAAAGCCCATGAT............. 807
        ||||      |||:::   :::|||||||:::   :::
    255 nGlyArgTyrGluIleMetAlaAspLysSerIleGluAlaAspTyrProG 272

    808 ...................GCCGCCGTGTTCAACCGCAATGCCGAATAT 837
                            ||||||   :::|||:::   :::|||
    272 lnLysAspGluThrAlaAlaAlaAlaTyrMetAsnLysPheValGluGlu 288

    838 TGGATACGCCGTTTTCCGACGCAGTATCTGTTTATGTACAACCGCTACAA 887
           |||   |||   |||   |||:::::::::::::::::::|||:::||
    289 IleIleLeuArgAlaProGluGlnTrpMetTrpLeuHisLysArgPheLy 305

    888 A 888
        |
    305 s 305
```

seq_name: sp_bacteriap:Q92P56

seq_documentation_block:

```
ID   Q92P56      PRELIMINARY;      PRT;   312 AA.
AC   Q92P56;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update).
DE   PUTATIVE LIPID A BIOSYNTHESIS KDO2-LAUROYL-LIPID IVA
DE   ACYLTRANSFERASE PROTEIN (EC 2.3.1.).
OS   Rhizobium meliloti (Sinorhizobium meliloti).
OC   Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC   Rhizobiaceae; Sinorhizobium.
OX   NCBI_TaxID=382;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=1021;
RX   MEDLINE=21368234; PubMed=11474104;
RA   Galibert F., Finan T.M., Long S.R., Puehler A., Abola P., Ampe F.,
RA   Barloy-Hubler F., Barnett M.J., Becker A., Boistard P., Bothe G.,
RA   Boutry M., Bowser L., Buhrmester J., Cadieu E., Capela D., Chain P.,
RA   Cowie A., Davis R.W., Dreano S., Federspiel N.A., Fisher R.F.,
RA   Gloux S., Godrie T., Goffeau A., Golding B., Gouzy J., Gurjal M.,
RA   Hernandez-Lucas I., Hong A., Huizar L., Hyman R.W., Jones T., Kahn D.,
RA   Kahn M.L., Kalman S., Keating D.H., Kiss E., Komp C., Lelaure V.,
RA   Masuy D., Palm C., Peck M.C., Pohl T.M., Portetelle D., Purnelle B.,
RA   Ramsperger U., Surzycki R., Thebault P., Vandenbol M.,
RA   Vorhoelter F.J., Weidner S., Wells D.H., Wong K., Yeh K.-C., Batut J.;
RT   "The composite genome of the legume symbiont Sinorhizobium meliloti.";
RL   Science 293:668-672(2001).
DR   EMBL; AL591789; CAC46516.1; -.
KW   Transferase; Acyltransferase; Complete proteome.
SQ   SEQUENCE   312 AA;  35743 MW;  06DF241E473D9985 CRC64;
```

alignment_scores:

```
            Quality:  135.50          Length:    303
              Ratio:  0.847             Gaps:     12
Percent Similarity:  52.805   Percent Identity: 23.102
```

alignment_block:
US-09-303-518D-569 x Q92P56 ..

Align seg 1/1 to: Q92P56 from: 1 to: 312

```
     43 GCCATGCACATCCTGTTGACCGCCCTGCTCAAATGCCTCTCCCTGCTGCC 92
        |||:::   :::||||||      ||||||      |||:::|||:::||
     24 AlaGlnPheValLeuLeu......LeuLeuThrIleLeuLysLeuPhePr 38

     93 GCTTTCCTGTCTGCACACGCTGGGAAACCGGCTCGGACATCTGGCGTTTT 142
        |   :::         ::::::   :::|||:::::::::: ||| :::
     38 oAlaAspMetAlaIleAsnPheMetAspArgValAlaArg.Trp.IleGl 54

    143 ACCTTTTAAAGGAAGACCGCGCGCGCATCGTCGCCAATATGCGTCAGGCA 192
         |||      ||||||   :::       :::   |||:::|||:::|||
     54 yProLysThrGlyArgHisLysLeu.ThrLeuIleAsnLeuArgAsnAla 70

    193 GGCATGAATCCCGACCCCAAAACGGTCAAAGCCGTTTTTGCGGAAACG.. 240
           :::   :::         ::::::|||:::     |||:::
     71 PheProGluLysSerAspAlaGluIleGluAlaIleAlaLeuGluSerTr 87

    241 ...................GCAAAAGGCGGTTTGGAACTTGCCCCCGCGT 271
                            |||         :::   |||         |
     87 pGlyAsnMetGlyArgLeuAlaAlaGluTyrValPheLeuAspGlnLeuP 104

    272 TTTTCAGAAAACCGGAAGACATAGAAACAATGTTCAAAGCGGTACACGGC 321
         ||         |||         |||                |||   |||
    104 heAspPheAspProAlaArgAlaGluProGlyArgValGluValSerGly 120

    322 TGGGAACATGTGCAGCAGGCTTTGGACAAACACGAAGGGCTGCTATTCAT 371
                       :::      |||:::   :::   ::::::   ::
    121 IleProLeuPheLeuGluLeuArgAspAsnProArgProPheIleValPh 137

    372 CACGCCGCACATCGGCAGCTACGATTTGGGCGGACGCTACATCAGCCAGC 421
        :|||   |||   |||:::::::::::::
    137 eThrAlaHisSerGlyAsnPheGluMet..................... 146

    422 AGCTTCCG.....................TTCCCGCTGACCGCCATGTAC 450
          ||||||                    :::   :::|||:::::::::
    147 ..LeuProValAlaGlySerAlaPheGlyLeuAspValThrValLeuPhe 162

    451 AAACCGCCGAAAATCAAAGCGATA...GACAAAATCATGCAGGCGGGCAG 497
        :::||||||:::          :::   |||:::::::::::::      ::
    163 ArgProProAsnAsnProTyrLeuAlaAspArgValPheGluPheArgLy 179

    498 GGTTCGCGGCAAAGGAAAAACCGCGCCTACCAGCATACAAGGGGTCAAAC 547
        :   |||      |||:::   :::|||:::         :::
    179 sGluArg...MetGlyAsnLeuValProSerHisAlaGlySerSerPheA 195

    548 AAATCATCAAAGCCCTGCGTTCGGGCGAAGCAACCATCGTCCTGCCCGAC 597
        :::   :::   |||:::   |||   ::::::   ||||||   |||
    195 laLeuAlaArgGlnLeuGluArgGlyGlyGlyValGlyValLeuValAsp 211

    598 CACGTCCCCTCCCCTCAAGAAGGCGGGGAAGGCGTATGGGTGGATTTCTT 647
                     |||:::      :::|||:::   :::   |||||
    212 ..............GlnLysPheArgLysGlyLeuLysThrLysPhePh 223

    648 CGGCAAACCTGCCTATACCATGACGCTGGCGGCAAAATTGGCACACGTCA 697
        |::::::   :::   |||      |||   |||||||||:::::
    223 eAlaGlnAspValGlnThrAsnProLeuLeuAlaLysLeuValArgGlnP 240

    698 AAGGCGTGAAAACCCTGTTTTTCTGCTGCGAACGCCTGCCTGGCGGACAA 747
        :::   ::::::            |||   ||||||||||||:::::::
    240 heAsnCysGluValTyrProAlaArgCysIleArgLeuProGlyAsnArg 256

    748 GGTTTCGATTTGCACATCCGCCCCGTC..............CAAGGGGA 782
          |||   |||:::::::::|||:::              ::::::::
    257 ...PheArgLeuGluLeuGluProAlaValGluIleProArgArgAlaAs 272
```

```
     783 ATTGAACGGCGACAAAGCCCATGATGCCGCCGTGTTCAACCGCAATGCCG 832
         :   :::   |||   :::      |||   ::::::|||   ::::::|
     272 pGlySerValAspValThrAlaThrAlaGlnMetLeuAsnAspLysValG 289

     833 AATATTGGATACGCCGTTTTCCGACGCAGTATCTGTTTATGTACAACCGC 882
         ||   |||:::|||::::::|||   |||:::|||:::   ::::::|||
     289 luSerTrpValArgGluTyrProGlyGlnTrpLeuTrpTyrHisAspArg 305

     883 TAC 885
         :::
     306 Trp 306

seq_name: sp_bacteria:Q9F6E3

seq_documentation_block:
ID   Q9F6E3      PRELIMINARY;      PRT;   320 AA.
AC   Q9F6E3;
DT   01-MAR-2001 (TrEMBLrel. 16, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE   PKSA PUTATIVE O-METHYLTRANSFERASE.
OS   Streptomyces collinus.
OC   Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC   Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces.
OX   NCBI_TaxID=42684;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=DSM2012;
RA   Martin R., Bailey J.E., Minas W.;
RT   "Putative spore pigment polyketide synthase cluster from Streptomyces
RT   collinus DSM2012.";
RL   Submitted (AUG-2000) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF293354; AAG26883.1; -.
KW   Transferase; Methyltransferase.
SQ   SEQUENCE   320 AA;  35453 MW;  AEDF4519923A166F CRC64;


alignment_scores:
           Quality:  135.00              Length:    320
             Ratio:  0.951                Gaps:     15
 Percent Similarity:  44.375   Percent Identity: 27.500

alignment_block:
 US-09-303-518D-569 x Q9F6E3   ..

 Align seg 1/1  to: Q9F6E3  from: 1  to: 320

      27 TCCCCCTTTGCGAACCGCCATGCACATCCTGTT.........GACCGCCC 67
         ||||||         |||:::   |||   :::      |||
      44 SerProValLeuProArgArgLeuHisGlyThrProGlyArgAspValAr 60

      68 TGCTCAAATGCCTCTCCCTGCTGCCGCTTTCCTGTCTGCACACGCTGGGA 117
         |||:::      ||||||||||||||||                 |||:
      60 gAlaHisArgArgLeuProAlaAlaAla..............ArgGlyA 72

     118 AACCGGCTCGGACATCT.............................. 134
         ::|||   :::   :::
      72 rgProProGlnProAlaArgHisHisAlaValValHisArgAlaValAsp 88

     135 ...GGCGTTTTACCTTTTAAAGGAAGACCGCGCGCGCATCGTCGCCAATA 181
            |||   |||         ||||||      |||||||||   |||
      89 LeuGlyArgLeuAlaGluAlaGlyArgGlyGlyAlaHisArgProGlnAr 105

     182 TGC.........GTCAGGCAGGCATGAATCCCGACCCCAAAACGGTCAA 221
                    |||||||||:::   :::|||   |||:::|||::::
     105 gGlyGlyGlyProValArgGlnGlyValLeuProLeuProGlnArgGlyA 122

     222 AGCCGTTTTTGCGGAAACGGCAAAAGGCGGTTTGGAACTTGCCCCCGCGT 271
         ::|||   :::      |||                 |||   ||||||
     122 rgProGlyValGlyGlyArg.................LeuGlnProArg 132

     272 TTTTCAGAAAACCGGAAGACATAGAAACAATGTTCAAAGCGGTACACGGC 321
            ::::::   ::: |||

     133 HisAspAspLeuGln.Arg............................... 138

     322 TGGGAACATGTGCAGCAGGCTTTGGACAAACACGAAGGGCTGCTATTCAT 371
                    |||:::|||   |||:::   ||||||:::
     139 ..........AlaValGlyAlaGlyArg...ArgArgValProArgPro 150

     372 CACGCCGCACAT............CGGCAGCTACGATTTGGGCGGACGCT 409
         |||:::   |||             |||:::   |||   |||
     151 HisGlyMetHisLeuGlyGlyArgArgArgArgArgProGlyAlaArgAr 167

     410 ACATCAGCCAGCAGCTTCCGTTCCC..................... 434
            :::||||||   :::||||||
     167 gGlyGluProAlaArgGlnValProGlyProAlaArgAsp***AlaArgP 184

     435 .............GCTGACCGCCATGTACAAACCGCCGAAAATCAAAGC 470
                       ::::::|||:::   :::||||||:::::::::
     184 roProProGlyGlyGlyGluArgArgProGluThrAlaHisGlyArgArg 200

     471 GATAGACAAAATCATGCAGGCGGGCAGGGTTCGCGGCAAAGGAAAAACCG 520
            |||       :::|||       :::   :::|||:::|||   :::
     201 ProArgGlyProArgAla......ArgArgAlaArgArgArgProAspGl 215

     521 CGCCTACCAGCATACAAGGGGTCAAACAAATCATCAAAGCCCTGCGTTCG 570
                    :::   |||         :::::::::   |||:::   |
     215 yArgProGlyGluGlyArgArgLeuArgAspGlnGluHisProGlyValG 232

     571 GGCGAAGCAACCATCGTCCT...GCCCGACCACGTCCCCTCCCCTCAAGA 617
         |||||   :::|||:::|||   |||       ||||||          |||
     232 lyArgArgGluHisHisProHisAlaAlaGlnArgProArgGlyArgArg 248

     618 AGGCGGGGAAGGCGTATGGGTGGATTTCTTCGGCAAACCTGCCTATACCA 667
            |||       |||   |||             :::      |||
     249 ProArgAspAlaArgGlyGlyHisArgGlu***ArgArgArgLeuAlaLe 265

     668 TGACGCTGGCGGCAAAATTGGCACACGTCAAAGGCGTGAAAACCCTGTTT 717
          ||||||                  |||:::|||:::::::|||
     265 uAspAla.................ValGlnHisArgAspGlyPro.... 274

     718 TTCTGCTGCGAACGCCTGCCTGGCGGACAAGGTTTCGATTTGCACATCCG 767
                    :::   |||   :::   |||         |||||||||
     275 ..........AlaAlaAlaProGlnArgArgArgGlyGlnAlaHisPro 287

     768 CCCCGTCCAAGGGGAATTGAACGGCGACAAAGCCCATGATGCCGCCGTGT 817
            :::      |||      :::|||      |||      ||||||
     288 AlaGluHisGlyGlyProAlaHisArgGlyArgProGlyHisArgArgHi 304

     818 TCAACCGCAATGCCGAATATTGGATACGCCGTTTTCCGACGCAGTATCTG 867
          ||||||                  |||:::|||      |||   :
     304 sGlnPro..............GlyGlnProLeuSerAlaArgValArgL 316

     868 TTTATGTA 875
         ::::::::
     316 euHisIle 318

seq_name: sp_bacteria:O06470

seq_documentation_block:
ID   O06470      PRELIMINARY;      PRT;   405 AA.
AC   O06470;
DT   01-JUL-1997 (TrEMBLrel. 04, Created)
DT   01-JUL-1997 (TrEMBLrel. 04, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   CEOA.
GN   CEOA.
OS   Burkholderia cepacia (Pseudomonas cepacia).
OC   Bacteria; Proteobacteria; beta subdivision; Burkholderia group;
OC   Burkholderia.
OX   NCBI_TaxID=292;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=K61-3;
RA   Charron R.L., Barry J.J., Burns J.L.;
RL   Submitted (APR-1997) to the EMBL/GenBank/DDBJ databases.
```

```
DR   EMBL; U97042; AAB58160.1; -.
DR   InterPro; IPR002215; HlyD.
DR   Pfam; PF00529; HlyD; 1.
SQ   SEQUENCE   405 AA;  44441 MW;  CCDC22ECFA35167F CRC64;


alignment_scores:
            Quality:  133.50              Length:    212
              Ratio:  1.362                Gaps:     10
 Percent Similarity:  46.226   Percent Identity: 28.774

alignment_block:
 US-09-303-518D-569 x O06470   ..

 Align seg 1/1  to: O06470  from: 1  to: 405

     150 AAAGGAAGACCGCGCGCGCATCGTCGCCAATATGCGTCAGGCAGGCATGA 199
         |||   :::|||:::::::::|||   :::   :::::::|||   |||::
     161 LysArgGluProGluGlyArgArgSerArgAlaGlyAsnGlyAlaHisGl 177

     200 ATCCCGACCCCAAAACGGTCAAAGCCGTTTTTGCGGAAACGGCAAAAGGC 249
         :||||||                                  :::   :::|
     177 nSerArgLeuTyr.........................AlaHisHisA 185

     250 GGTTTGGAACTTGCCCCCGCGTTTTTCAGAAAACCGGAAGACATAGAAAC 299
         ||   |||:::   |||||||||   :::      |||:::|||   :::
     185 rgAlaGlyValGlyProArgValAlaArg......GlyAsnHisAlaArg 199

     300 AATGTTCAAAGCGGTACACGGCTGGGAACATGTGCA......GCAGGCTT 343
         :::   :::   |||   |||:::|||   :::|||      ||||||::
     200 GlnArgArgValGlyArgArgValGlyArgAlaAlaAspAspAlaGlyIl 216

     344 TGGA.....................CAAACACGAAGGGCT.........G 363
         :|||                     :::   |||||||||         |
     216 eGlyValAlaAspLeuArgValValArgArgArgArgAlaAspLeuProA 233

     364 CTATTCATCAC..................................... 374
         |||||||||:::
     233 laIleHisGlnArgArgAlaGlnArgProGlnGlyAlaGlyArgAlaArg 249

     375 .....................GCCGCACATCGGCAGCTACGATTTGGGCG 403
                              |||:::|||:::   ||||||||||
     250 ProArgGluArgAsnArgLeuLeuAlaGlnArgArgAspArgPheGlyAr 266

     404 GACGCTACATCAGCCAGCAGCTTCCGTTCCCGCTGACCGCCATGTACAAA 453
         |         |||||||||      |||      ||||||
     266 g.........GlnProAlaArgHisValVal......ArgHisAspProA 278

     454 CCGCCGAAAATCAAAGCGATAGACAAAATCATGCAGGC............ 491
         |||:::         :::|||   :::|||   |||
     278 rgAlaArgProLeuArgGlnArgGlyArgHisProGlyProGlyProLeu 294

     492 ......GGGCAGGGTTCGCGGCAAAGGAAAAACCGCGCCTACCAGCATAC 535
               :::|||:::||||||||||      |||:::      :::
     295 ArgThrArgGluGlyGlyArgGlnArgAlaAlaArgGlyAlaAlaArgAr 311

     536 AAGGGGTCAAACAAATCATCAAAGCCCTGCGTTCGGGCGAAGCAACCATC 585
         |||         :::|||:::                |||   ::::::|
     311 gArgArgGlyAspGlnHisArgProGlyGlnGluValArgValArgArgA 328

     586 GTCCTGCCCGACCACGTCCCCTC.......................... 608
         ||||||||   ||||||   :::
     328 rgProAlaGlyProArgValValSerArgSerAlaAlaArgAspAlaAla 344

     609 ......CCCTCAAGAAGGCGGGGAAGGCGTATGGGT 638
               |||:::|||   |||   |||   .:::|||
     345 ArgGlnProAlaArgAspArgGluArgAlaValGly 356

 seq_name: sp_bacteria:Q51415

 seq_documentation_block:
 ID   Q51415      PRELIMINARY;      PRT;   684 AA.
 AC   Q51415;
 DT   01-NOV-1996 (TrEMBLrel. 01, Created)
 DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
 DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
 DE   ALGINATE LYASE.
 GN   ALGY.
 OS   Pseudomonas aeruginosa.
 OC   Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
 OC   Pseudomonas.
 OX   NCBI_TaxID=287;
 RN   [1]
 RP   SEQUENCE FROM N.A.
 RC   STRAIN=A22 ALG+;
 RA   Schmitt Andrieu L., Hulen C.;
 RT   "Cloning and sequencing of a new alginate lyase, algY, from mucoid
 RT   strains of Pseudomonas aeruginosa, and its expression in E. coli.";
 RL   Submitted (SEP-1995) to the EMBL/GenBank/DDBJ databases.
 DR   EMBL; Z54213; CAA90940.1; -.
 KW   Lyase.
 SQ   SEQUENCE   684 AA;  74958 MW;  CBCCE638F5675D7A CRC64;


 alignment_scores:
             Quality:  129.50              Length:    254
               Ratio:  1.028                Gaps:     10
  Percent Similarity:  49.606   Percent Identity: 26.378

 alignment_block:
  US-09-303-518D-569 x Q51415   ..

  Align seg 1/1  to: Q51415  from: 1  to: 684

       42 CGCCATGCACATCCTGTTGAC...CGCCCTGCTCAAATGCCTCTCCCTGC 88
          |||:::   ||||||   :::      |||:::              |||::
      304 ArgArgArgHisProPheSerAlaThrProSerAlaThrArgSerProTh 320

       89 TGCCGCTTTCCTGTCTGCACACGCTGGGAAACCGGCTCGGACATCTGGCG 138
          :::::::      |||:::   |||::::::||||||      :::|||
      320 rSerSer......SerGlyLeuAlaAlaArgProAlaGlyArgGlyGlyG 335

      139 TTTTACCTTTTAAAGGAAGACCGCGCGCGCATCGTCGCCAATATGCGTCA 188
            ||||||   :::::::::|||||:::::|||:::         |||
      335 lyLeuProArgGlnAlaHisProArgGlyGluArgGlnGlyAlaLeuSer 351

      189 GGCAGGCATGAATCCCGACCCCAAAACGGTCAAAGCCGTTTTTGCGGAAA 238
          :::      :::                          ||||||   |||
      352 AlaAlaAlaThrAlaValSerAlaAlaValLeuArgArgPheLeuGly.. 367

      239 CGGCAAAAGGCGGTTTGGAACTTGCCCCCGCGTTTTTCAGAAAACCGGAA 288
                                                       ||||
      368 .............................................GlyA 369

      289 GACATAGAAACAATGTTCAAAGCGGTACACGGCTG............GGA 326
          ||   |||:::::::   :::::::|||   ||||||
      369 rgProArgThrGlyArgArgThrGlyGlyArgLeuProAspLeuGlyArg 385

      327 ACATGTGCAGCAGGCTTTGGACAAACACGAAGGGCTGCTATTCATCACGC 376
          :::   :::::::|||   |||:::::||||||   |||         ||
      386 AlaValGlyGlyGlyGlyGlyGluAsnArgArgArgAla......ProAl 400

      377 CGCACATCGGCAGCTACGATTTGGGCGGACGCTACATCAGCCAGCAGCTT 426
          |:::                 :::|||::::::   |||   ||||||:::
      400 aSerGlySerProTrpAlaTyrGlyGluValArgHisProProAlaCysA 417

      427 CCGTTCCCGCTGACCGCCATGTACAAACCGCC...GAAAATCAAAGCGAT 473
          ::   |||   :::|||:::   :::   |||   ::::::|||:::|||
      417 spArgProArgAsnArgArgArgSerLeuAlaArgArgGlyGlnAlaAsp 433

      474 AGACAAAATCATGCAGGCGGGCAGGGTTCGCGGCAAAGGAAAAACCGCGC 523
          |||   :::         ::::::      |||                |||
      434 ArgThrHisLeuArgArgAsnHisArgArgArgAlaAlaValLeuArgPr 450

      524 CTACCAGCATACAAGGGGTCAAACAAATCATCAAAGCCCTGCGTTCGGGC 573
           ::::::      |||      |||:::                  ||||
```

```
      450 oPheArgSerGlyArgProAlaThrHis.................GlyA 461

      574 GAAGCAACCATCGTCCTGCCCGACCACGTCCCCTCCCCTCAAGAAGGCGG 623
          ||:::   :::||||||   |||           |||::::::
      461 rgAlaAlaArgArgProProArgSerProGlyAspProAlaGluProLeu 477

      624 GGAAGGCGTATGGGTGGATTTCTTCGGCAAACCTGCCTATACCATGACGC 673
          ||||||         |||      |||:::           |||   ||
      478 GlyArgGlyArgProGlyAlaArgArgArgArg.........HisArgAl 491

      674 TGGCGGCAAAATTGGCACA.....................CGTCAAA 699
          ||||   :::   ::::::                     |||:::
      491 aGlyArgArgSerAlaAlaGlyGlyGlyAlaHisArgArgIleArgArgA 508

      700 GGCGTGAAAACCCTGTTTTTCTGCTGCGAACGCCTGCCTGGCGGACAAGG 749
            |||::::::   :::   |||||||||      |||    |||:::||
      508 laArgHisArgGlnLeuHisLeuLeuArgLeuSerAla.ProGlyArgGl 524

      750 TTTCGATTTG 759
          |:::   |||
      524 yLeuProLeu 527

seq_name: sp_human:Q9UQ39

seq_documentation_block:
ID   Q9UQ39      PRELIMINARY;      PRT;   956 AA.
AC   Q9UQ39;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   RNA BINDING PROTEIN (FRAGMENT).
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Ohtaki S., Umeki K., Sawada Y.;
RT   "Homo sapiens mRNA for RNA binding protein, partial cds.";
RL   Submitted (JUL-1998) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AB016088; BAA83714.1; -.
DR   InterPro; IPR002965; P_rich_extensn.
DR   PRINTS; PR01217; PRICHEXTENSN.
FT   NON_TER       1      1
FT   NON_TER     956    956
SQ   SEQUENCE   956 AA;  106462 MW;  C9FEAD07B49F5011 CRC64;


alignment_scores:
           Quality:  128.50            Length:    317
             Ratio:  0.905              Gaps:     18
 Percent Similarity:  44.795   Percent Identity: 26.814

alignment_block:
 US-09-303-518D-569 x Q9UQ39   ..

 Align seg 1/1  to: Q9UQ39  from: 1  to: 956

       27 TCCCCCTTTGCGAACCGCCATGCACATCCTGTTGACCGCCCTGCTCAAAT 76
          ||||||      :::|||   :::          :::   ||||||
      316 SerProSerProGluArgSerSerThrGlyProGluProProAlaProTh 332

       77 GCCTCTCCCTGCTGCCGCTTTCCTGTCTGCACACGCTGGGAAACCG.... 122
           ||||||   |||            :::   |||:::|||:::|||
      332 rProLeuLeuAla...........GluArgHisGlyGlySerProGlnP 345

      123 .....GCTCGGACA..........................TCTGGC 137
               |||   |||                         |||
      345 roLeuAlaThrThrProLeuSerGlnGluProValAsnProProSerGlu 361

      138 GTTTTACCTTTTAAAGGAAGA......CCGCGCGCGCATCGTCGCCAATA 181
          :::   |||   :::   |||      |||::::::   ::::::
      362 AlaSerProThrArgAspArgSerProProLysSerProGluLysLeuPr 378

      182 TGCGTCAGGCAGGCATGAATCCCGACCCCAAAACGGTCAAAGCCGTTTTT 231
          |||:::      ||||||   |||   :::   |||
      378 oGlnSerSerSerSerGluSerSerProProSerProGlnProThrLysV 395

      232 GCGGAAACGGCAAAAGGCGGTTTGGAACTTGCCCCCGCGTTTTTCAGAAA 281
            ::::::                          |||:::
      395 alSerArgHisAlaSerSerSerProGluSerProLysProAlaProAla 411

      282 ACCGGAAGACATAGAAACAATGTTCAAAGCGGTACACGGCTGGGAACATG 331
            |||   |||||||:::      :::|||   |||      :::
      412 ProGlySerHisArgGluIleSerSerSerProThrSerLysAsn..... 426

      332 TGCAGCAGGCTTTGGACAAACACGAAGGGCTGCTATTCAT...CACGCCG 378
                :::   |||:::::::::|||          |||   |||:::
      427 ....ArgSerHisGlyArgAlaLysArgAspLysSerHisSerHisThrP 442

      379 CACATCGGCAGCTACGATTTGGGCGGACGCTACATCAGCCAGCAGCTTCC 428
            |||       |||:::||||||:::   ::::::||||||::::::
      442 roSerArg......ArgMetGlyArgSer...ArgSerProAlaThrAla 455

      429 .........................GTTCCCGCTGACCGC...CATGTACA 451
                                  :::|||:::   |||   |||   ::
      456 LysArgGlyArgSerArgSerArgThrProThrLysArgGlyHisSerAr 472

      452 AACCGCCGAAAATCAAAGCGATAGACAAAATCATGCA..... 488
          ::::   :::   |||      |||::::::   |||
      472 gSerArgSerProGlnTrpArgArgSerArgSerAlaGlnArgTrpGlyA 489

      488 ........................................... 488

      489 rgSerArgSerProGlnArgArgGlyArgSerArgSerProGlnArgPro 505

      489 GGCGGGCAGGGTTCGCGG......CAAAGGAAAAACCGCGCCTACCAGCA 532
          |||   :::   ||||||      ||||||::::::|||:::   :::
      506 GlyTrpSerArgSerArgAsnThrGlnArgArgGlyArgSerArgSerAl 522

      533 TACAAGGGGTCAAACAAATCATCAAAGCCCTGCGTTC...GGGCGAAGCA 579
            ||||||:::::::::   :::|||||||||      |||||||||:
      522 aArgArgGlyArgSerHisSerArgSerProAlaThrArgGlyArgSerA 539

      580 ACCATCGT...CCTGCCCGA............CCACGTCCCCTCCCCTCA 614
          ::   |||   |||||||||                        |||:::
      539 rgSerArgThrProAlaArgArgGlyArgSerArgSerArgThrProAla 555

      615 AGAAGGCGGGGAAGGCGTATGGGTGGATTTCTTCGGCAAACCTGCCTATA 664
          ||||||||||:::|||
      556 ArgArgArgSerArg.................................. 560

      665 CCATGACGCTGGCGGCAAAATTGGCACACGTCAAAGGCGTGAAAACCCTG 714
                    ::::::      ||||||:::|||   ::::::
      561 .............SerArgThrProThrArgArgArgSerArgSer.... 571

      715 TTTTTCTGCTGCGAACGCCTGCCTGGCGGACAAGGTTTCGATTTGCACAT 764
                    ||||||||||||   |||   |||   |||
      572 ...........ArgThrProAlaArgArgGlyArgSerArgSerArgThr 584

      765 CCGCCCCGTCCAAGGGGAATTGAACGGCGACAAAGCCCATGATGCCGCCG 814
          |||   |||   |||      :::|||   :::||||||      |||||
      585 ProAlaArgArgArg......SerArgThrArgSerProValArgArgAr 599

      815 T 815
          |
      599 g 599

seq_name: sp_human:Q9UQ40

seq_documentation_block:
ID   Q9UQ40      PRELIMINARY;      PRT;  1262 AA.
AC   Q9UQ40;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
```

```
DE   RNA BINDING PROTEIN (FRAGMENT).
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Ohtaki S., Umeki K., Sawada Y.;
RT   "Homo sapiens mRNA for RNA binding protein, partial cds.";
RL   Submitted (JUL-1998) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AB016087; BAA83713.1; -.
DR   InterPro; IPR002965; P_rich_extensn.
DR   PRINTS; PR01217; PRICHEXTENSN.
FT   NON_TER       1      1
FT   NON_TER    1262   1262
SQ   SEQUENCE   1262 AA;  138003 MW;  75E76BD73FA14295 CRC64;


alignment_scores:
            Quality:  128.50            Length:    317
              Ratio:  0.905               Gaps:     18
 Percent Similarity:  44.795   Percent Identity: 26.814

alignment_block:
 US-09-303-518D-569 x Q9UQ40   ..

 Align seg 1/1  to: Q9UQ40  from: 1  to: 1262

      27 TCCCCCTTTGCGAACCGCCATGCACATCCTGTTGACCGCCCTGCTCAAAT 76
         ||||||     :::|||  :::       :::   ||||||
     141 SerProSerProGluArgSerSerThrGlyProGluProProAlaProTh 157

      77 GCCTCTCCCTGCTGCCGCTTTCCTGTCTGCACACGCTGGGAAACCG.... 122
          ||||||   |||          :::   |||:::|||:::|||
     157 rProLeuLeuAla.....GluArgHisGlyGlySerProGlnP 170

     123 .....GCTCGGACA.........................TCTGGC 137
              |||  |||                          |||
     170 roLeuAlaThrThrProLeuSerGlnGluProValAsnProProSerGlu 186

     138 GTTTTACCTTTTAAAGGAAGA......CCGCGCGCGCATCGTCGCCAATA 181
         :::   |||   :::   |||      |||::::::   ::::::
     187 AlaSerProThrArgAspArgSerProProLysSerProGluLysLeuPr 203

     182 TGCGTCAGGCAGGCATGAATCCCGACCCCAAAACGGTCAAAGCCGTTTTT 231
          |||:::      ||||||  |||   :::   |||
     203 oGlnSerSerSerSerGluSerSerProProSerProGlnProThrLysV 220

     232 GCGGAAACGGCAAAAGGCGGTTTGGAACTTGCCCCCGCGTTTTTCAGAAA 281
         ::::::                          |||:::
     220 alSerArgHisAlaSerSerSerProGluSerProLysProAlaProAla 236

     282 ACCGGAAGACATAGAAACAATGTTCAAAGCGGTACACGGCTGGGAACATG 331
         |||   ||||||:::      :::|||   |||      :::
     237 ProGlySerHisArgGluIleSerSerSerProThrSerLysAsn..... 251

     332 TGCAGCAGGCTTTGGACAAACACGAAGGGCTGCTATTCAT...CACGCCG 378
             :::   |||::::::::::|||         |||   |||:::
     252 ....ArgSerHisGlyArgAlaLysArgAspLysSerHisSerHisThrP 267

     379 CACATCGGCAGCTACGATTTGGGCGGACGCTACATCAGCCAGCAGCTTCC 428
          |||      |||:::||||||:::   ::::::||||||::::::
     267 roSerArg......ArgMetGlyArgSer...ArgSerProAlaThrAla 280

     429 .......................GTTCCCGCTGACCGC...CATGTACA 451
                                :::|||:::   |||   |||   ::
     281 LysArgGlyArgSerArgSerArgThrProThrLysArgGlyHisSerAr 297

     452 AACCGCCGAAAATCAAAGCGATAGACAAAATCATGCA............. 488
         ::::   :::   |||      |||::::::   |||
     297 gSerArgSerProGlnTrpArgArgSerArgSerAlaGlnArgTrpGlyA 314

     488 ..................................................... 488

     314 rgSerArgSerProGlnArgArgGlyArgSerArgSerProGlnArgPro 330

     489 GGCGGGCAGGGTTCGCGG......CAAAGGAAAAACCGCGCCTACCAGCA 532
         |||   :::   ||||||      ||||||::::::|||:::   :::
     331 GlyTrpSerArgSerArgAsnThrGlnArgArgGlyArgSerArgSerAl 347

     533 TACAAGGGGTCAAACAAATCATCAAAGCCCTGCGTTC...GGGCGAAGCA 579
             ||||||::::::::::   :::|||||||||      |||||||||:
     347 aArgArgGlyArgSerHisSerArgSerProAlaThrArgGlyArgSerA 364

     580 ACCATCGT...CCTGCCCGA............CCACGTCCCCTCCCCTCA 614
         ::   |||   |||||||||                          |||:::
     364 rgSerArgThrProAlaArgArgGlyArgSerArgSerArgThrProAla 380

     615 AGAAGGCGGGGAAGGCGTATGGGTGGATTTCTTCGGCAAACCTGCCTATA 664
         ||||||||||:::|||
     381 ArgArgArgSerArg..................................... 385

     665 CCATGACGCTGGCGGCAAAATTGGCACACGTCAAAGGCGTGAAAACCCTG 714
                     ::::::      ||||||::::||   ::::::
     386 .............SerArgThrProThrArgArgArgSerArgSer.... 396

     715 TTTTTCTGCTGCGAACGCCTGCCTGGCGGACAAGGTTTCGATTTGCACAT 764
                     ||||||||||||   |||   |||   |||
     397 ...........ArgThrProAlaArgArgGlyArgSerArgSerArgThr 409

     765 CCGCCCCGTCCAAGGGGAATTGAACGGCGACAAAGCCCATGATGCCGCCG 814
         |||   |||   |||       :::|||   :::||||||      |||||
     410 ProAlaArgArgArg......SerArgThrArgSerProValArgArgAr 424

     815 T 815
         |
     424 g 424

seq_name: sp_human:Q9UHA8

seq_documentation_block:
ID   Q9UHA8      PRELIMINARY;      PRT;  2296 AA.
AC   Q9UHA8;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   SPLICING COACTIVATOR SUBUNIT SRM300.
GN   SRM300.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=20132238; PubMed=10668804;
RA   Blencowe B.J., Bauren G., Eldridge A.G., Issner R., Nickerson J.A.,
RA   Rosonina E., Sharp P.A.;
RT   "The SRm160/300 splicing coactivator subunits.";
RL   RNA 6:111-120(2000).
DR   EMBL; AF201422; AAF21439.1; -.
DR   InterPro; IPR002965; P_rich_extensn.
DR   PRINTS; PR01217; PRICHEXTENSN.
SQ   SEQUENCE   2296 AA;  251964 MW;  17C0BD4EA10A9CF9 CRC64;


alignment_scores:
            Quality:  128.50            Length:    317
              Ratio:  0.905               Gaps:     18
 Percent Similarity:  44.795   Percent Identity: 26.814

alignment_block:
 US-09-303-518D-569 x Q9UHA8   ..

 Align seg 1/1  to: Q9UHA8  from: 1  to: 2296

      27 TCCCCCTTTGCGAACCGCCATGCACATCCTGTTGACCGCCCTGCTCAAAT 76
```

```
     |||||||      :::|||   :::          :::   |||||||
 351 SerProSerProGluArgSerSerThrGlyProGluProProAlaProTh 367

  77 GCCTCTCCCTGCTGCCGCTTTCCTGTCTGCACACGCTGGGAAACCG.... 122
      |||||||   |||            :::   |||:::|||:::|||
 367 rProLeuLeuAla............GluArgHisGlyGlySerProGlnP 380

 123 .....GCTCGGACA............................TCTGGC 137
           |||   |||                           |||
 380 roLeuAlaThrThrProLeuSerGlnGluProValAsnProProSerGlu 396

 138 GTTTTACCTTTTAAAGGAAGA......CCGCGCGCGCATCGTCGCCAATA 181
     :::   |||   :::   |||      |||::::::   ::::::
 397 AlaSerProThrArgAspArgSerProProLysSerProGluLysLeuPr 413

 182 TGCGTCAGGCAGGCATGAATCCCGACCCCAAAACGGTCAAAGCCGTTTTT 231
         |||:::      ||||||   |||   :::   |||
 413 oGlnSerSerSerSerGluSerSerProProSerProGlnProThrLysV 430

 232 GCGGAAACGGCAAAAGGCGGTTTGGAACTTGCCCCCGCGTTTTTCAGAAA 281
       ::::::                        |||:::
 430 alSerArgHisAlaSerSerSerProGluSerProLysProAlaProAla 446

 282 ACCGGAAGACATAGAAACAATGTTCAAAGCGGTACACGGCTGGGAACATG 331
     |||   |||||:::      :::|||   |||      :::
 447 ProGlySerHisArgGluIleSerSerSerProThrSerLysAsn..... 461

 332 TGCAGCAGGCTTTGGACAAACACGAAGGGCTGCTATTCAT...CACGCCG 378
           :::   |||:::::::::|||         |||   |||:::
 462 ....ArgSerHisGlyArgAlaLysArgAspLysSerHisSerHisThrP 477

 379 CACATCGGCAGCTACGATTTGGGCGGACGCTACATCAGCCAGCAGCTTCC 428
       |||      |||:::|||||||:::   ::::::||||||::::::
 477 roSerArg......ArgMetGlyArgSer...ArgSerProAlaThrAla 490

 429 .......................GTTCCCGCTGACCGC...CATGTACA 451
                            :::|||:::   |||   |||   ::
 491 LysArgGlyArgSerArgSerArgThrProThrLysArgGlyHisSerAr 507

 452 AACCGCCGAAAATCAAAGCGATAGACAAAATCATGCA.............. 488
     ::::   :::   |||      |||::::::   |||
 507 gSerArgSerProGlnTrpArgArgSerArgSerAlaGlnArgTrpGlyA 524

 488 ................................................... 488

 524 rgSerArgSerProGlnArgArgGlyArgSerArgSerProGlnArgPro 540

 489 GGCGGGCAGGGTTCGCGG......CAAAGGAAAAACCGCGCCTACCAGCA 532
     |||   ::   ||||||      ||||||:::::::|||:::   :::
 541 GlyTrpSerArgSerArgAsnThrGlnArgArgGlyArgSerArgSerAl 557

 533 TACAAGGGGTCAAACAAATCATCAAAGCCCTGCGTTC...GGGCGAAGCA 579
       ||||||::::::::   :::|||||||||      |||||||||:
 557 aArgArgGlyArgSerHisSerArgSerProAlaThrArgGlyArgSerA 574

 580 ACCATCGT...CCTGCCCGA............CCACGTCCCCTCCCCTCA 614
     ::   |||   |||||||||                    |||:::
 574 rgSerArgThrProAlaArgArgGlyArgSerArgSerArgThrProAla 590

 615 AGAAGGCGGGGAAGGCGTATGGGTGGATTTCTTCGGCAAACCTGCCTATA 664
     |||||||||:::|||
 591 ArgArgArgSerArg.................................. 595

 665 CCATGACGCTGGCGGCAAAATTGGCACACGTCAAAGGCGTGAAAACCCTG 714
               ::::::      ||||||:::|||   ::::::
 596 .............SerArgThrProThrArgArgArgSerArgSer.... 606

 715 TTTTTCTGCTGCGAACGCCTGCCTGGCGGACAAGGTTTCGATTTGCACAT 764
             ||||||||||||   |||   |||   |||
 607 ...........ArgThrProAlaArgArgGlyArgSerArgSerArgThr 619

 765 CCGCCCCGTCCAAGGGGAATTGAACGGCGACAAAGCCCATGATGCCGCCG 814
     |||   |||   |||      :::|||   :::|||||||      |||||
 620 ProAlaArgArgArg......SerArgThrArgSerProValArgArgAr 634

 815 T 815
       |
 634 g 634
```

seq_name: sp_human:Q9UQ35

seq_documentation_block:

```
ID   Q9UQ35      PRELIMINARY;      PRT;  2752 AA.
AC   Q9UQ35;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   RNA BINDING PROTEIN.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Ohtaki S., Umeki K., Sawada Y.;
RT   "Homo sapiens mRNA for RNA binding protein, complete cds.";
RL   Submitted (JUL-1998) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AB016092; BAA83718.1; -.
DR   InterPro; IPR002965; P_rich_extensn.
DR   PRINTS; PR01217; PRICHEXTENSN.
SQ   SEQUENCE   2752 AA;  299672 MW;  109C64F181097123 CRC64;
```

alignment_scores:

```
           Quality:  128.50              Length:    317
             Ratio:  0.905                 Gaps:     18
 Percent Similarity:  44.795   Percent Identity: 26.814
```

alignment_block:

US-09-303-518D-569 x Q9UQ35 ..

Align seg 1/1 to: Q9UQ35 from: 1 to: 2752

```
  27 TCCCCCTTTGCGAACCGCCATGCACATCCTGTTGACCGCCCTGCTCAAAT 76
     |||||||      :::|||   :::          :::   |||||||
 351 SerProSerProGluArgSerSerThrGlyProGluProProAlaProTh 367

  77 GCCTCTCCCTGCTGCCGCTTTCCTGTCTGCACACGCTGGGAAACCG.... 122
      |||||||   |||            :::   |||:::|||:::|||
 367 rProLeuLeuAla............GluArgHisGlyGlySerProGlnP 380

 123 .....GCTCGGACA............................TCTGGC 137
           |||   |||                           |||
 380 roLeuAlaThrThrProLeuSerGlnGluProValAsnProProSerGlu 396

 138 GTTTTACCTTTTAAAGGAAGA......CCGCGCGCGCATCGTCGCCAATA 181
     :::   |||   :::   |||      |||::::::   ::::::
 397 AlaSerProThrArgAspArgSerProProLysSerProGluLysLeuPr 413

 182 TGCGTCAGGCAGGCATGAATCCCGACCCCAAAACGGTCAAAGCCGTTTTT 231
         |||:::      ||||||   |||   :::   |||
 413 oGlnSerSerSerSerGluSerSerProProSerProGlnProThrLysV 430

 232 GCGGAAACGGCAAAAGGCGGTTTGGAACTTGCCCCCGCGTTTTTCAGAAA 281
       ::::::                        |||:::
 430 alSerArgHisAlaSerSerSerProGluSerProLysProAlaProAla 446

 282 ACCGGAAGACATAGAAACAATGTTCAAAGCGGTACACGGCTGGGAACATG 331
     |||   |||||:::      :::|||   |||      :::
 447 ProGlySerHisArgGluIleSerSerSerProThrSerLysAsn..... 461

 332 TGCAGCAGGCTTTGGACAAACACGAAGGGCTGCTATTCAT...CACGCCG 378
           :::   |||:::::::::|||         |||   |||:::
 462 ....ArgSerHisGlyArgAlaLysArgAspLysSerHisSerHisThrP 477

 379 CACATCGGCAGCTACGATTTGGGCGGACGCTACATCAGCCAGCAGCTTCC 428
```

```
          |||      |||:::|||||||:::   :::::::||||||::::::
    477 roSerArg......ArgMetGlyArgSer...ArgSerProAlaThrAla 490

    429 .......................GTTCCCGCTGACCGC...CATGTACA 451
                               :::|||:::   |||   |||   :: .
    491 LysArgGlyArgSerArgSerArgThrProThrLysArgGlyHisSerAr 507

    452 AACCGCCGAAAATCAAAGCGATAGACAAAATCATGCA..... 488
        ::::   :::   |||      |||::::::   |||
    507 gSerArgSerProGlnTrpArgArgSerArgSerAlaGlnArgTrpGlyA 524

    488 ..................................................... 488

    524 rgSerArgSerProGlnArgArgGlyArgSerArgSerProGlnArgPro 540

    489 GGCGGGCAGGGTTCGCGG......CAAAGGAAAAACCGCGCCTACCAGCA 532
        |||   :::   ||||||      ||||||::::::|||:::   :::
    541 GlyTrpSerArgSerArgAsnThrGlnArgArgGlyArgSerArgSerAl 557

    533 TACAAGGGGTCAAACAAATCATCAAAGCCCTGCGTTC...GGGCGAAGCA 579
            ||||||:::::::::   :::||||||||||      |||||||||:
    557 aArgArgGlyArgSerHisSerArgSerProAlaThrArgGlyArgSerA 574

    580 ACCATCGT...CCTGCCCGA............CCACGTCCCCTCCCCTCA 614
        ::   |||   |||||||||                          |||:::
    574 rgSerArgThrProAlaArgArgGlyArgSerArgSerArgThrProAla 590

    615 AGAAGGCGGGGAAGGCGTATGGGTGGATTTCTTCGGCAAACCTGCCTATA 664
        ||||||||||:::|||
    591 ArgArgArgSerArg..................................... 595

    665 CCATGACGCTGGCGGCAAAATTGGCACACGTCAAAGGCGTGAAAACCCTG 714
                      ::::::      ||||||:::|||   ::::::
    596 .............SerArgThrProThrArgArgArgSerArgSer.... 606

    715 TTTTTCTGCTGCGAACGCCTGCCTGGCGGACAAGGTTTCGATTTGCACAT 764
                  ||||||||||||   |||   |||   |||
    607 ...........ArgThrProAlaArgArgGlyArgSerArgSerArgThr 619

    765 CCGCCCCGTCCAAGGGGAATTGAACGGCGACAAAGCCCATGATGCCGCCG 814
        |||   |||   |||      :::|||   :::||||||      |||||
    620 ProAlaArgArgArg......SerArgThrArgSerProValArgArgAr 634

    815 T 815
        |
    634 g 634
```

```
seq_name: sp_plant:Q9ARV4

seq_documentation_block:
ID   Q9ARV4      PRELIMINARY;      PRT;   395 AA.
AC   Q9ARV4;
DT   01-JUN-2001 (TrEMBLrel. 17, Created)
DT   01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   P0503E05.24 PROTEIN.
GN   P0503E05.24.
OS   Oryza sativa (Rice).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC   Ehrhartoideae; Oryzeae; Oryza.
OX   NCBI_TaxID=4530;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=CV. NIPPONBARE;
RA   Sasaki T., Matsumoto T., Yamamoto K.;
RT   "Oryza sativa nipponbare(GA3) genomic DNA, chromosome 1, PAC
RT   clone:P0503E05.";
RL   Submitted (DEC-2000) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AP003021; BAB40020.1; -.
SQ   SEQUENCE   395 AA;  42387 MW;  57DBEE7AA401412C CRC64;
```

```
alignment_scores:
            Quality:  128.00              Length:    316
              Ratio:  0.948                Gaps:     15
 Percent Similarity:  42.722   Percent Identity: 26.582

alignment_block:
 US-09-303-518D-569 x Q9ARV4   ..

 Align seg 1/1  to: Q9ARV4  from: 1  to: 395

      36 GCGAACCGCCATGCACATCCTGTTGACCGCCCTGCTCAAATGCCTCTCCC 85
         :::|||||||||            :::   |||         ||||||
      65 SerAsnArgHisLysIleGluArgSerSerProProTrpSerProLeuLy 81

      86 TGCTGCCGCTTTCCTGTCTGCACACGCTGGGAAACCGGCTCGGACATCTG 135

      81 s..................................................  81

     136 GCGTTTTACCTTTTAAAGGAAGACCGCGCGCGCATCGTCGCCAATATGCG 185
              |||           ||||||   ||||||:::
      82 ........ProThrAspHisPheProArgArgHisArgGlnProAspPhe 95

     186 TCAGGCAGGCATGAATCCCGACCCCAAAACGGTCAAAGCCGTTTTTGCGG 235
         :::|||   :::      |||   |||:::|||            |||||
      96 AlaGlyLeuArgAlaArgArgGlyGlnArgGly.....CysGl 108

     236 AAACGGC......AAAAGGCGGTTTGGA.....ACTTGCCCCC 267
         |:::|||      :::||||||   |||            :::|||   :
     108 yAspGlyAspAspArgArgArgGlyGlyGlyGlyArgSerCysGluG 125

     268 GCGTTTTTCAGAAAACCGGAAGACATAGAAACAATGTTCAAAGCGGTACA 317
         :::::   ::::::      |||   |||:::::::|||:::::::
     125 lnAlaAlaGluGluGlyArgArgGlyArgSerSerValArgAlaProPro 141

     318 CGGCTGGGAACA.............TGTGCAGCAGGCTTTGGACAAAC 352
             :::|||                :::||||||||||   |||   ||
     142 LeuAlaAlaThrHisLysAlaHisGluAlaAlaAlaGlyGlyGlyAlaTh 158

     353 A.........CGAAGGGCTGCTAT.......................... 367
         |         :::|||||||||
     158 rGlyArgArgGlnArgAlaAlaGlyGlyProGlyGlyGlyGlyArgSerA 175

     368 ........TCATCACGCCGCACA............TCGGCAGCTACGATT 397
                 :::|||||||||             :::   ||||||
     175 spGlyArgCysThrArgArgThrValLeuProGluAspArgAlaThrThr 191

     398 TGGGCGGACGCTACATCAGCCAGCAGCTTCCGTTCCCGCTGACCGCCATG 447
         |||:::   |||::::::   :::|||   |||::: |||   :::||||
     192 TrpGly...AlaAlaGlyArgGlnSerArgArgGlu.AlaGlyLysHisV 207

     448 TACAAACCGCCGAAAATCAAAGCGATAGACAAAATCATGCAGGCGGGCAG 497
         ||         |||:::   :::|||   ::::::   :::      :::
     207 alAlaGluGluGluAspGlyAlaAspGlyGluArgAspGlyGluCysGlu 223

     498 ......................GGTTCGCGGCAAAGGAAAAACCGCGC 523
                                |||   |||:::|||   :::
     224 AspAlaAlaAlaValValValHisGlyArgArgSerArgAlaSerAlaLe 240

     524 CTACCAGCATACAAGGGGTCAAACAAATCATCAAAGCCCTGCGTTCGGGC 573
                 :::   |||:::::::::      :::   |||   ||||
     240 uAsnGlyValAlaGlyGlySerValThrAlaThrGlyAspAlaValGlyA 257

     574 GAAGCAACCATCGTCCTGCCCGA.......................... 596
         ||       ||||||:::|||
     257 rgArgAlaGlyArgProValArgValAlaGluGlyGlyValValThrGlu 273

     597 ...CCACGTCCCCTCCCCTCAAGAAGGCGGGGAAGGCGTATGGGTGGATT 643
             ||||||   |||      |||:::|||   |||      |||
     274 LeuAspArgProSerProPheSerArgGlnGlyThrArgLysArgGlyGl 290

     644 TCTTCGGCAAACCTGCCTATACCATGACGCTGGCGGCAAAATTGGCACAC 693
         |||                  |||:::||||||:::         |
     290 yArgArgCysArgGlyArgGlyAspAspGlyGlyGlyGln.........A 304
```

```
     694 GTCAAAGGCGTGAAAACCCTGTTTTTCTGCTG.............CGA 728
         ||||||||   ::::::      :::||||||              |||
     304 rgGlnArgProHisArgArgSerTrpLeuLeuTrpTrpGlnArgValArg 320

     729 ACGCCTGCCTGGCGGACAAGGTTTCGATTTGCACAT.............. 764
             |||   |||:::   :::      ||||||
     321 GlyThrAlaAlaArgAlaProLeuLeuProAlaHisProThrGlnLeuLe 337

     765 .CCGCCCCGTCCAAGGGGAATTGAACGGCGACAAAGCCCATGATGC 809
          ||||||   |||          :::   |||         |||
     337 uProProProProMetProAlaAlaHisAlaGlnLeuAlaSerCys 352

seq_name: sp_bacteria:Q9ZI82

seq_documentation_block:
ID   Q9ZI82      PRELIMINARY;      PRT;   323 AA.
AC   Q9ZI82;
DT   01-MAY-1999 (TrEMBLrel. 10, Created)
DT   01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT   01-MAY-1999 (TrEMBLrel. 10, Last annotation update)
DE   LIPID A ACYLTRANSFERASE.
GN   MSBB.
OS   Salmonella typhimurium.
OC   Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC   Salmonella.
OX   NCBI_TaxID=602;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=ATCC14028;
RX   MEDLINE=99116994; PubMed=9920266;
RA   Low K.B., Ittensohn M., Le T., Platt J., Sodi S., Amoss M., Ash O.,
RA   Carmichael E., Chakraborty A., Fisher J., Lin S.L., Luo X., Miller S.,
RA   Zheng L.-M., King I., Pawelek J., Bermudes D.;
RT   "Lipid A mutant Salmonella with suppressed virulence and TNFalpha
RT   induction retain tumor-targeting in vivo.";
RL   Nat. Biotechnol. 17:37-41(1999).
DR   EMBL; AF039020; AAD03801.1; -.
KW   Transferase; Acyltransferase.
SQ   SEQUENCE   323 AA;  37287 MW;  3E77B3F35E997C47 CRC64;


alignment_scores:
            Quality:  127.50            Length:    290
              Ratio:  0.885               Gaps:     10
 Percent Similarity:  49.655   Percent Identity: 24.138

alignment_block:
 US-09-303-518D-569 x Q9ZI82   ..

 Align seg 1/1  to: Q9ZI82  from: 1  to: 323

      64 GCCCTGCTCAAATGCCTCTCCCTGCTGCCGCTTTCCTGTCTGCACACGCT 113
         |||   :::      ::::::|||   |||   |||             ||
      31 AlaAlaMetAlaGlyIleAlaLeuThrProAlaSerPheArgAspProLe 47

     114 GGGAAACCGGCTCGGACATCTGGCGTTTTACCTTTTAAAGGAAGACCGCG 163
         |         ||||||:::::::|||      |||   |||:::::::|||
      47 uLeuAlaThrLeuGlyArgPheAlaGlyArgLeuGlyLysSerSerArgA 64

     164 CGCGCATCGTCGCCAATATGCGT............CAGGCAGGCATGAAT 201
         |||   :::   |||:::                 |||   :::
      64 rgArgAlaLeuIleAsnLeuSerLeuCysPheProGlnArgSerGluAla 80

     202 CCCGACCCCAAAACGGTCAAAGCCGTTTTTGCGGAAACGGCAAAAGGCGG 251
                    |||      :::|||||||   :::   ::::::
      81 GluArgGluAlaIleValAspGluMetPheAlaThrAlaProGlnAlaMe 97

     252 TTTGGAACTTGCCCCCGCGTTTTTCAGAAAACCGGAAGACATAGAAACAA 301
               :::|||         :::|||   |||:::   |||:::
      97 tAlaMetMetAlaGluLeuAlaMetArgGlyProLysLysIleGlnGlnA 114

     302 TGTTCAAAGCGGTACACGGCTGGGAACATGTGCAGCAGGCTTTGGACAAA 351
                      :::|||   |||   :::::::::         :::
     114 rgValAsp...TrpGluGlyLeuGluIleIleGluGluMetArgArgAsn 129

     352 CACGAAGGGCTGCTATTCATCACGCCGCACATCGGCAGCTACGATTTGGG 401
         |||   :::::::|||:::::::||||||      :::   |||:::
     130 AspGluLysValIlePheLeuValProHisGlyTrpGlyValAspIlePr 146

     402 CGGACGCTACATCAGCCAGCAGCTTCCGTTCCCGCTGACCGCCATGTACA 451
          :::      :::::::::|||           ::::::||||||:::
     146 oAlaMetLeuMetAlaSerGln...GlyGlnLysMetAlaAlaMetPheH 162

     452 AACCGCCGAAAATCAAAGCGATAGACAAAATCATGCAGGCGGGCAGGGTT 501
                        :::::::|||   |||   ::::::   |||
     162 isAsnGlnGlyAsnProValPheAspTyrIleTrpAsnThrValArgArg 178

     502 CGCGGCAAAGGAAAAACCGCGCCTACCAGCATACAAGGGGTCAAACAAAT 551
         |||      |||:::             :::   :::|||:::|||   ::
     179 ArgPheGlyGlyArgLeuHisAlaArgAsn...AspGlyIleLysProPh 194

     552 CATCAAAGCCCTGCGTTCGGGCGAAGCAACCATCGTCCTGCCC......G 595
         :|||::::::::::|||:::|||                 ||||||      |
     194 eIleGlnSerValArgGlnGlyTyrTrpGlyTyrTyrLeuProAspGlnA 211

     596 ACCACGTCCCCTCCCCTCAAGAAGGCGGGGAAGGCGTATGGGTGGATTTC 645
         |||||   |||:::   :::|||                 :::|||||||||
     211 spHisGlyProGluHisSerGlu.............PheValAspPhe 222

     646 TTCGGCAAACCTGCCTATACCATGACGCTGGCGGCAAAATTGGCACACGT 695
         |||:::            |||:::         :::::|||      ||
     223 PheAlaThrTyrLysAlaThrLeuProAlaIleGlyArgLeuMetLysVa 239

     696 CAAAGGCGTGAAAACCCTGTTTTTCTGCTGCGAACGCCTG.......... 735
         |                      |||      |||:::
     239 l.....................CysArgAlaArgValIleProLeuP 248

     736 ..CCTGGCGGACAAGGT............TTCGATTTGCACATCCGCCCC 771
           |||      :::|||            :::   ::::::|||||||||
     248 heProValTyrAsnGlyLysThrHisArgLeuThrIleGlnIleArgPro 264

     772 GTCCAAGGGGAA...TTGAACGGCGACAAAGCCCATGATGCCGCCGTGTT 818
         :::   :::   |||:::::|||            |||      ::
     265 ProMetAspAspLeuLeuThrAlaAspAspHisThrIleAlaArgArgMe 281

     819 CAACCGCAATGCCGAATATTGGATACGCCGTTTTCCGACGCAGTATCTGT 868
         :|||:::::::::|||   :::::          |||   ||||||   :
     281 tAsnGluGluValGluIlePheValGlyProHisProGluGlnTyrThrT 298

     869 TTATGTACAACCGCTACAAA 888
         :::::   :::      |||
     298 rpIleLeuLysLeuLeuLys 304

seq_name: sp_human:Q16824

seq_documentation_block:
ID   Q16824      PRELIMINARY;      PRT;   797 AA.
AC   Q16824;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   PROFILAGGRIN (FRAGMENT).
GN   FLG.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=91064347; PubMed=2248957;
RA   Gan S.Q., McBride O.W., Idler W.W., Markova N., Steinert P.M.;
RT   "Organization, structure, and polymorphisms of the human profilaggrin
RT   gene [published erratum appears in Biochemistry 1991 Jun
RT   11;30(23):5814].";
RL   Biochemistry 29:9432-9440(1990).
```

```
DR   EMBL; M60502; AAA63248.1; -.
DR   InterPro; IPR003303; Filaggrin.
DR   PRINTS; PR00487; FILAGGRIN.
FT   NON_TER      1       1
SQ   SEQUENCE   797 AA;  85176 MW;  60E6184763BDA86B CRC64;


alignment_scores:
            Quality:  125.50            Length:    222
              Ratio:  1.037               Gaps:     10
 Percent Similarity:  54.505   Percent Identity: 24.324

alignment_block:
 US-09-303-518D-569 x Q16824   ..

 Align seg 1/1  to: Q16824  from: 1  to: 797

     162 CGCGCGCATCGTCGCCAATATGCGTCAGGCAGGCATGAATCCCGACCCCA 211
         |||:::   |||||||||   :::|||      :::::::|||      ||
     266 ArgThrSerArgArgGlnGlySerServalSerGlnAspSerAspSerGl 282

     212 AAACGGTCAAAGCCGTTTTTGCGGAAACGGCAAAAGGCGGTTTGGAACTT 261
         |   |||:::|||:::         :::::::::||||||   |||:::: 
     282 n...GlyHisSerGlu.........AspSerGluArgArgSerGlySerA 295

     262 GCCCCCGCGTTTTTCAGAAAACCGGAAGACATAGAAACAATGTTCAAAGC 311
         ::   |||                 :::||||||::::::::|||::: 
     295 laSerArg.................AsnHisArgGlySerAlaGlnGlu 305

     312 GGTACACGGCTGGGAACA.....................TGTGCAGC 337
           :::|||   |||:::                      ||
     306 GlnSerArgAspGlySerArgHisProArgSerHisHisGluAspArgAl 322

     338 AGGCTTTGGA............CAAACACGAAGGGCTGCTATTCATCACG 375
         ||||   |||           ::::::|||:::::::::   |||||||
     322 aGlyHisGlyAspSerAlaGluSerSerArgGlnSerGlyThrHisHisA 339

     376 CCGCACATCGGCAGCTACGATTTGGGCGGACGCTACATCAGCCAGCAGCT 425
         ||   :::   :::         :::   :::   |||:::   |||
     339 laGluAsnSerSerGlyGlyGlnAlaAlaSerSerHisGluGlnAlaArg 355

     426 TCCGTTCCCGCTGACCGCCATGTACAAACCGCCGAAAATCAAAGCGATAG 475
         |||      ::::::|||||   :::      ::::::::::|||
     356 SerSerAlaGlyGluArgHisGlySerHisTyrGlnGlnSerAlaAspSe 372

     476 ACAAAATCATGCAGGC.........GGGCAGGGTTCGCGGCAAAGGAAAA 516
          ::::::|||:::|||         |||||||:::|||       ::::
     372 rSerArgHisSerGlyIleGlyHisGlyGlnAlaSerSerAlaValArgA 389

     517 ACCGCGCCTACCAGCATACAAGGGGT..................... 542
         ::   :::::::::   :::   |||
     389 spSerGlyHisArgGlySerSerGlySerGlnAlaSerAspAsnGluGly 405

     543 .........CAAACAAATCATCAAAGCCCTGCGTTCGGGCGA......AG 577
                 :::::::::   ||||||   :::      |||      ::
     406 HisSerGluAspSerAspThrGlnSerValSerAlaHisArgGlnAlaGl 422

     578 CAACCATCGTCCTGCCCGACCACGTCCCCTCCCCTCAAGAAGGCGGGGAA 627
         ::::|||:::   ::::::   :::         :::|||   |||||||
     422 yArgHisHisGluSerHisGlnGluSerThrArgGlyArgSerArgGlyA 439

     628 GG.....CGTATGGGTGGATTTCTTCGGCAAACCTGCCTATACCATGAC 671
         ||      |||   |||:::||||||   |||:::      :::::::::
     439 rgSerGlyArgSerGlySerPheLeuTyrGlnValSerThrHisGluGln 455

     672 GCT..............................GGCGGCAAAATTGG 688
         :::                             |||||||:::   ||
     456 SerGluSerAlaHisGlyArgAlaGlyProSerThrGlyGlyArgGlnGl 472

     689 CACACGTCAAAGGCGT 704
         |:::|||:::::::::
     472 ySerArgHisGluGln 477
```

```
seq_name: sp_bacteria:Q52747

seq_documentation_block:
ID   Q52747      PRELIMINARY;      PRT;   680 AA.
AC   Q52747;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   ENDO-GLUCANASE.
OS   Ruminococcus flavefaciens.
OC   Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridiaceae;
OC   Ruminococcus.
OX   NCBI_TaxID=1265;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=186;
RA   Huang C.M., Asmundson R.V., Yu P.L.;
RT   "Nucleotide sequence of a cellulase gene complex from Ruminococcus
RT   flavefaciens strain 186 coding for multi-cellulase activities.";
RL   Submitted (SEP-1990) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; X56082; CAA39559.1; -.
SQ   SEQUENCE   680 AA;  75503 MW;  5A589F334846DD6F CRC64;


alignment_scores:
            Quality:  125.00            Length:    317
              Ratio:  0.919               Gaps:     21
 Percent Similarity:  42.902   Percent Identity: 28.076

alignment_block:
 US-09-303-518D-569 x Q52747   ..

 Align seg 1/1  to: Q52747  from: 1  to: 680

      29 CCCCTTTGCGAACCGCCATGCACATCCTGTTGACCG...CCCTGCTCAAA 75
         |||           |||   |||:::      |||   |||   |||
      27 ProThrProProSerProSerThrGluSerSerProThrProSerSerPr 43

      76 TGCCTCTCCC.............TGCTGCCGC.........TTTCCTG 101
          |||||||||                 ||||||            |||:
      43 oAlaSerProArgSerProThrValArgCysArgGluAlaValAlaProA 60

     102 TCTGCACACGCTGGGAAACCGGCTCGGACATCTGGCGTTTTACCTTTTAA 151
         ::||| ||||||         |||||||||:::|||            ::
      60 laCys.HisAla......AspAlaArgThrAlaGly............Ar 70

     152 AGGAAGA...CCGCGCGCGCATCGTCGCCAATATGCGTCAGGCAGGCATG 198
         :||||||   ||||||   :::                |||||||||
      70 gGlyArgAspProArgHisGlnCysAspAlaArgLeuArgGlyArgHis. 86

     199 AATCCCGACCCCAAAACGGTCAAAGCCGTTTTTGCGGAAACGGCAAAAGG 248
               |||||||::   |||::::::              :::|||
      87 .....ArgProArgLeuGlyHisAlaValValArgAspAspGly...... 99

     249 CGGTTTGGAACTTGCCCCCGCGTTTTTCAGAAAACCGGAAGACATAGAAA 298
                                                |||   |||:::::
     100 ..............................LeuGlyProHisGlnAr 105

     299 CAATGTTCAAAGCGGTACACGGCTGGGAACATGTGCA......GCAGGCT 342
         ::::         |||:::      |||      :::      |||   :
     105 gAspCysValArgGlyAlaCysHisGlyProSerValHisArgAlaAspT 122

     343 TTGGACAAACACGAAGGGCTGCTATTCATCACGCCGCACATCGGCAGCTA 392
         ::   |||   ||||||||||||:::::
     122 rpHisGlnProArgArgAlaAlaLeuGlu..................... 131

     393 CGATTTGGGCGGACGCTACATCAGCCAGCAGCTTCCGTTCCCGCTGACCG 442
               |||:::   :::|||      ::::::   ||||||:::   ||
     132 LeuValGlyGlnLeuValHisProGlnSerGlyLeuValProValValAr 148

     443 CCATGTA.................CAAACCGCCGAAAATCAAAGC.... 470
         |   |||                 :::      ::::::||||||
     148 gGlyValProAlaProProProArgAspArgArgArgAspGlnSerArgP 165
```

```
      471 ..................................GATAGACAAAAT... 482
                                           ||||||:::::
      165 roArgArgAlaLeuValGlnLeuProLeuProArgAspArgGluGlnPro 181

      483 ..............................CATGCAGGCGGGCAGGG 499
                                       |||:::||||||:::||
      182 GlyGlyHisGlnProHisAspAlaAspAlaValHisGlyGlyGlyArgGl 198

      500 TTCGCGG...CAAAGGAAAAACCGCGCCTACCAGCATACAAGGGGTCAAA 546
          |   |||   |||   ::::::|||      :::      |||   :::
      198 yValArgGlyGlnLeuGlnSerArgArgProArgAspArgArgValArgA 215

      547 CAAATCATCAAAGCCCTGCGTTCGGGCGAAGCAACCATCGTCCTGCCCGA 596
            :::::::::::::|||:::   |||||||||         |||   |||
      215 spHisArgAspGlyProSerAspGlyArgSerPheAspProProAspArg 231

      597 CCACGTCCCCTCCCCTCAAGAAGGCGGGGAAGGCGTATGGGTGGATTTCT 646
            |||   :::      |||:::||||||||||||
      232 GlnArgAlaVal...ArgArgGlnArgGlyArgArg.............. 242

      647 TCGGCAAACCTGCCTATACCATGACGCTGGCGGCAAAATTGGCACACGTC 696
                          |||                 :::     |||
      243 ..................HisArgLeu.........ValAspArgArgV 250

      697 AAAGG...CGTGAAAACCCTGTTTTTCTGCTGCGAACGCCTGCCTGGCGG 743
            |||   ||||||:::||||||     .:::|||      :::
      250 alArgProArgGluGlnProValProArgValArgGlnArgGlyProGly 266

      744 ACAAGGTTTCGATTTGCACATCCGCCCCGTCCAAGGGGA........... 782
          :::|||         ||||||      |||   |||
      267 ValArg.........AlaHisGluHisArgArgArgArgGlyProLeuAl 280

      783 ..........ATTGAACGGCGACAAAGCCCATGATGCCGCCGTGTTCAA 821
                   :::|||||||   |||:::         |||   ||||||
      280 aGlyAspLeuValGluArgAlaGlnAlaAlaHisProArgSerValGln 296
```

seq_name: sp_bacteria:Q53478

seq_documentation_block:

```
ID   Q53478      PRELIMINARY;      PRT;   456 AA.
AC   Q53478;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   ORF2 PROTEIN.
GN   ORF2.
OS   Neisseria gonorrhoeae.
OC   Bacteria; Proteobacteria; beta subdivision; Neisseriaceae; Neisseria.
OX   NCBI_TaxID=485;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=95020543; PubMed=7934834;
RA   Facius D., Meyer T.F.;
RT   "A novel determinant (comA) essential for natural transformation
RT   competence in Neisseria gonorrhoeae and the effect of a comA defect on
RT   pilin variation.";
RL   Mol. Microbiol. 10:699-712(1993).
DR   EMBL; S75490; AAB32262.1; -.
SQ   SEQUENCE   456 AA;  54380 MW;  30AC7B5CDFB912D5 CRC64;
```

alignment_scores:

```
            Quality:  124.00              Length:    343
              Ratio:  0.800                Gaps:     26
 Percent Similarity:  45.190   Percent Identity: 25.656
```

alignment_block:

```
 US-09-303-518D-569 x Q53478   ..

 Align seg 1/1  to: Q53478  from: 1  to: 456

      32 CTTTGCGAACCGCCATGCACATCCTGTTGACCGCCCTGCTCAAATGCCTC 81
```

```
          :::|||:::||| |||   |||            |||   ::::::|||
     143 ValCysArgPro.HisLysHisArgGlnAlaLeuPro...ArgIleProA 158

      82 TCCCTGCTGCCGCTTTCCTGTCTGCACACGCTGGGAAACCGGCTCGGACA 131
                 :::|||:::      ::::::            ||||||
     158 snArgAsnArgSerPheValPheArgTyrCys.........PheArgThr 171

     132 TCTGGCGTTTTACCTTTT......AAAGGAAGACCGCGC........... 164
             |||      |||         :::   |||||||||
     172 ValGlyTyrLysProCysGlnGlnGlnHisArgProArgThrGlnThrGl 188

     165 ....GCGCATCGTCGCCAATATGCGTCAGGCAGGCATGAATCCCGACCCC 210
             |||::::::::::|||      ||||||   |||:::::::::   |
     188 nAlaAlaArgGlnGlnGlnGlnGlnSerGlyThrHisArgGlnHisGlnG 205

     211 AAAACGGTCAAAGCCGTTTTTGCGGAAACGGCAAA......AGGCGGTTT 254
         ||      :::::::|||      ||||||||||||      |||::::::
     205 ln......ArgGlnArgCysGlyGlyAsnGlyLysLeuArgArgHisIle 219

     255 GGA.....................................ACTTGCCCCCG 268
         |||                                    |||||||||
     220 GlyLysProHisGlnLysProGlnSerValPheArgGlnLysCysProAr 236

     269 CGTTTTTCAGAAAACCGGAAGACATAGAAACAATGTTCAAAGCGGTACAC 318
         |::::::::::::      ||||||   :::::::::|||:::      |
     236 gLeuLeuLysArgArgLysArgHis...AspArgThrGlnGlnArgGlnA 252

     319 GGCTGGGAACAT..........................GTGCAGCAGGC 341
         ||    :::|||                          ::::::|||
     252 rg.GlyGlnHisProArgLysProArgTyrArgHisArgIleAspGlnAr 268

     342 TTTGGACAAACACGAAGGGCTGCTATTCATCACGCCGCACATCGGCAGCT 391
             |||:::::::::
     268 gAlaAspGlnArgGln.................................. 273

     392 ACGATTTGGGCGGACGCTACATCAGCCAGCAG.................. 423
                   |||:::   |||::::::|||
     274 ........ArgGlyLysGlnIleThrGluGlnAsnGlnArgGlnHisArg 287

     424 ..........................CTTCCGTTCCCGCTG......AC 440
                                   ||||||||||||||      ||
     288 SerArgProLeuProAlaHisGlyGlyLeuProPheProLeuValGlnTh 304

     441 CGCCATGTACA.........AACCGCCGAAAATCAAAGCGATAGACAAAA 481
         | |||||||||          :::   ||| :::|||::::::||||||
     304 r.ProCysThrArgProTyrGlnGlyArg.HisGlnThrGluArgGlnPr 320

     482 TCATGCAGGCGGGCAGGGTTCGCGGCAAAGGAAAAACCGCGCCTACCAGC 531
             :::      |||   :::||||||||||   ::::::      ||||
     320 oSerThrAspThrGln...AsnArgGlnArgValGluGln......GlnH 334

     532 ATACAAGGGGTCAAACAAATCATCAAAGCCCTGCGTTCGGGCGAAGCAAC 581
         ||   |||   :::          :::   |||
     334 isArgArgGlnArgLeuProProArgArgPro.................. 344

     582 CATCGTCCTGCCCGACCACGTCCCCTCCCCTCAAGAAGGCGGGGAAGGCG 631
         :::|||   :::   |||   |||   |||   :::||||||   :::::
     345 ArgArgGlnSerSerProAlaProArgProArgGluArgArgGlnHisGl 361

     632 TATGGGTGGATTTCTTCGGCAAACCTGCCTATACCATGACGCTGGCGGCA 681
         :       ::::::   |||               |||        ||||
     361 nAsnThrAlaLeuArgArg..............HisGlyLysThrGlyL 373

     682 AAATTGGCACACGTCAA..............AGGCGTGAAAACCCTGTT 716
         ||      ::::::|||              ||||||:::|||||||
     373 ysGlnArgValLysGlnArgAlaArgThrProArgArgGlnAsnPro... 388

     717 TTTCTGCTGCGAACGCCTGCCTGGCGGACAAGGTTTCGATTTGCACATCC 766
                   |||:::   :::::::|||            :::||
     389 ......ArgAlaArgProGlyGlyLysAlaArgGlyThrGlyLysGlnPr 403

     767 G.........CCCCGTCCA..............AGGGGAATTGAACGGC 792
         |         ||||||                 |||:::::::::   |
```

```
    403 oLeuArgGlnProArgLysGlnHisArgHisHisArgAsnValGlnThrA 420

    793 GACAAAGCCCATGATGCCGCCGTGTTCAA..................... 821
        ||:::::         |||:::   |||
    420 rgAsnThrAspGlnMetArgGlnProGlnArgProGluArgLeuProGln 436

    822 ......CCGCAATGCCGAATA 836
              |||||||||:::
    437 SerGlyAlaGlnCysArgLeu 443

seq_name: sp_bacteria:Q9LCV5

seq_documentation_block:
ID   Q9LCV5      PRELIMINARY;      PRT;   886 AA.
AC   Q9LCV5;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   ATP OPERON (FRAGMENT).
OS   Rhodospirillum rubrum.
OC   Bacteria; Proteobacteria; alpha subdivision; Rhodospirillaceae;
OC   Rhodospirillum.
OX   NCBI_TaxID=1085;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=85251588; PubMed=2861810;
RA   Falk G., Hampe A., Walker J.E.;
RT   "Nucleotide sequence of the Rhodospirillum rubrum atp operon.";
RL   Biochem. J. 228:391-407(1985).
DR   EMBL; X02499; CAB97255.1; -.
DR   InterPro; IPR001899; Gram_pos_anchor.
DR   PROSITE; PS00343; GRAM_POS_ANCHORING; UNKNOWN_1.
FT   NON_TER     886    886
SQ   SEQUENCE   886 AA;  95954 MW;  BAE140D8802E6F6C CRC64;


alignment_scores:
            Quality:  123.50            Length:    433
              Ratio:  0.787               Gaps:     21
 Percent Similarity:  36.259   Percent Identity: 21.478

alignment_block:
 US-09-303-518D-569 x Q9LCV5   ..

 Align seg 1/1  to: Q9LCV5  from: 1  to: 886

      51 CATCCTGTTGACCGC..................................... 65
         ||||||:::   |||
      21 HisProThrGlyArgThrGlyLeuGlyAspArgLeuLeuIleAlaAlaMe 37

      66 ....CCTGCTCAAATGCCTCTCCCTGCTGCCGCTTTCCTGTCTGCACACG 111
               ||||||   |||:::   |||:::            :::   |
      37 tValThrAlaGlnAlaProIleAlaAlaValPheAsnGlnProSerGlyA 54

     112 CTGGGAAACCGGCTCGGACATCTGGCGTTTTACCTTTTAAAGGAAGACCG 161
         ||:::::::      :::          :::   |||   :::|||:::
      54 laAlaArgAlaPheGlnProMetThrAlaAlaProThrGlnGlyGlnArg 70

     162 CGC.....................................GCGCATCGTCG 175
         |||                                   |||:::|||::
      71 ArgIleAlaProAlaIleGluLysGlnGlnAsnLeuLeuAlaArgArgGl 87

     176 CCAATATGCGTCAGGCAGGCATGAATCCCGACCCCAAAACGGTCAAAGCC 225
         :   :::   :::|||      :::   |||   |||   :::|||:::
      87 nGlyPheLeuAspGlyGlyAspGlnArgArgAlaGlnProAlaGlnAlaG 104

     226 GTTTTTGCGGAAACGGCAAAAGGCGGTTTGGAACTTGCCCCCGCGTTTTT 275
           :::|||:::|||:::||||||
     104 lyArgAlaGlyHisGlyArgArgArgGlyArgArg.............. 115

     276 CAGAAAACCGGAAGACATAGAAACAAT..................... 302
         ::::::   ||||||:::|||::::::
     116 ArgArgGlyGlyArgArgArgArgGlyGlyArgPheLeuLeuLeuGlyLe 132

     303 ............GTTCAAAGCGGTACACGGCTGGGAACATGTGCAGCAG 339
                     :::      |||      ||||||         ||||
     132 uValValArgProAlaGlyArgGlyGlnThrLeuGly......ProAlaG 147

     340 GCTTTGGACAAACACGA.............................. 356
         ||   |||:::   |||
     147 lySerGlyLysIleArgGlyGlnProPheLeuArgArgGlyAlaAlaGln 163

     357 ........................AGGGCTGCTATTCATCACGCCGC 379
                                 |||:::|||:::      :::
     164 IleGluGlnGlyAspLeuGlyHisPheArgSerAlaValAlaThrGlyGl 180

     380 ACATCGG.....................................CAGC 390
          |||:::                                     |||
     180 nHisGlnMetProIleAlaAlaLeuAlaGlyValLeuAspGlyPheGlnA 197

     391 TACGATTTGGGCGGACGCTACAT...CAGCCAGCAGCTTCCGTTCCCGCT 437
           |||   ||||||      |||         ||||||:::::   |||
     197 rgArgArgGlyArgGlyGluHisGlyGlyLysAlaAlaGluAlaGlyAla 213

     438 GACCGC...CATGTACAAACCGCCGAAAATCAAAGC.............. 470
         |||:::   ||||||      :::   :::::::::
     214 AspHisGlyHisValAlaGlyValIleArgHisAlaLeuLeuLeuPheGl 230

     471 .....................GATAGACAAAATCATGCAGGCGGGCAGG 498
                              |||   ::::::      |||   |
     230 uGlyAlaValValPhePheIleAspAspAspGlnAlaGlnIleGlyIleG 247

     499 GTTCGCGGCAAAGGAAAAACCGCGCCTACCAGCATACAAGGGGTCAAACA 548
         ||::::::||||||::::::|||   ::::::|||   |||
     247 lyGlnLysGlnArgArgSerArgProHisHisHisArgArg......... 260

     549 AATCATCAAAGCCCTGCGTTCGGGCGAAGCAACCATCGTCCTGCCCGACC 598
                    |||      |||         |||||||||      ||
     261 ............ProProGlnGly.........HisArgProProSerPr 270

     599 ACGTCCC.....................CTCCCCTCAAGAAGGCGGGGA. 626
         |   |||                     :::|||   :::|||   |||
     270 oProProArgArgArgAlaHisLeuArgMetProLeuHisArgSerGlyA 287

     627 .......................AGGCGTATGGGTGGA............ 641
                               |||   :::||||||
     287 laGluAlaValGlyGluAlaProArgProLeuGlyGlyGlnGlyAspPhe 303

     641 ............................................... 641

     304 GlyGlnGlnHisGlnSerLeuProProAlaAlaGlnGlyPheGlyAspAr 320

     642 .....................................TTTCTTCGGC 651
                                              |||   :::|
     320 gLeuGluIleAspLeuGlyLeuAlaArgAlaGlyAspProPheGlnGlnG 337

     652 AAACCTGCCTATACCATGACGCTGGCGGCAAA.................. 683
         |||||                  ||||||
     337 lnThr.....................GlyLysSerLeuValLysArgArg 346

     684 ..........................ATTGGCACACGTCAAAGGCGTGA 706
                                  ||||||   ||||||:::|||::
     347 AlaGlnAlaValArgArgArgLeuLeuIleGly...ArgGlnGlnArgSe 362

     707 AAACCCTGTTTTTCTG......CTGCGAACGCCTGCCTGGCGGACAAGGT 750
         :::::||||||   :::      ||||||   |||      ||||||
     362 rGlyProValGlyIleGlyGlyLeuArgArgProPheIleArgThrGlyH 379

     751 TTCGATTTGCACATCCG.......................CCCCGTCCA 776
           |||:::   :::|||                        ||||||
     379 isArgLeuGluGlnProGlyLeuGlyHisArgArgHisHisArgArgPro 395

     777 AGGGGA.......................................... 782
         :::|||
     396 HisGlyArgPheProArgGlnIleGlyAlaGlyGluGlyGlnThrGlyMe 412
```

```
     783 ....ATTGAACGGCGACAAAGCCCATGATGCCGCCGTGTTCAACCGCAA 827
             :::::::::|||||:::|||   :::      :::|||||||:::
     412 tValValLysAsnArgGlnAsnProLeuAlaAlaGlyAlaGlnProArg 428

seq_name: sp_bacteria:Q9RA26

seq_documentation_block:
ID   Q9RA26      PRELIMINARY;      PRT;   313 AA.
AC   Q9RA26;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   GENES, SIMILAR TO EICOSAPENTAENOIC ACID SYNTHESIS GENE CLUSTER.
OS   Vibrio marinus (Moritella marina).
OC   Bacteria; Proteobacteria; gamma subdivision; Alteromonadaceae;
OC   Moritella.
OX   NCBI_TaxID=90736;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=MP-1;
RA   Tanaka M., Ueno A., Kawasaki K., Yumoto I., Ohgiya S., Hoshino T.,
RA   Ishizaki K., Okuyama H., Morita N.;
RT   "Isolation of clustered genes that are notably homologous to the
RT   eicosapentaenoic acid biosynthesis gene cluster from the
RT   docosahexaenoic acid-producing bacterium Vibrio marinus strain MP-1.";
RL   Biotechnol. Lett. 21:939-945(1999).
DR   EMBL; AB025342; BAA89377.1; -.
SQ   SEQUENCE   313 AA;  35840 MW;  86A5D6DAC3154ACE CRC64;


alignment_scores:
            Quality:  122.00            Length:    189
              Ratio:  1.089               Gaps:      6
 Percent Similarity:  59.259   Percent Identity: 24.868

alignment_block:
 US-09-303-518D-569 x Q9RA26   ..

 Align seg 1/1  to: Q9RA26  from: 1  to: 313

     313 GTACACGGCTGGGAACATGTGCAGCAGGCTTTGGACAAACACGAAGGGCT 362
         |||||||||   |||::::::         |||:::::::   ::::::
     117 ValHisGlyGlyGluAsnLeuProLeuLeuGluGlnGlyLysAlaCy 133

     363 GCTATTCATCACGCCGCACATCGGCAGCTACGATTTGGGCGGACGCTACA 412
          ::::::::::::|||||       :::   |||::::::|||   :::|
     133 sIleLeuLeuValProHisSerPheAlaIleAspPheAlaGlyLeuHisI 150

     413 TCAGCCAGCAGCTTCCGTTCCCGCTGACCGCCATGTACAAACCGCCGAAA 462
         ||:::   :::          |||:::   :::|||::::::      :::
     150 leAla...SerTyrGlyAlaProPheCysThrMetPheAsnAsnSerGlu 165

     463 ATCAAAGCGATAGACAAAATCATGCAGGCGGGCAGGGTTCGCGGCAAAGG 512
         :::   :::|||   :::|||         |||:::         ||
     166 AsnGluLeuPheAspTrpLeuMetThrArgGlnArgAlaMet...PheGl 181

     513 AAAAACCGCGCCTACCAGCATACAAGGGGTCAAACAAATCATCAAAGCCC 562
         |   |||:::               |||:::      ::::::|||:::|
     181 yGlyThrValTyrHisArgLysAlaGlyLeuGlyAlaLeuValLysSerL 198

     563 TGCGTTCGGGCGAAGCAACCATCGTCCTGCCCGACCACGTCCCCTCCCCT 612
         ||:::|||||||||:::         |||||||||:::      :::|||
     198 euLysSerGlyGluSerCysTyrTyrLeuProAspGluAspHisGlyPro 214

     613 CAAGAAGGCGGGGAAGGCGTATGGGTGGATTTCTTCGGCAAACCTGCCTA 662
         :::         :::::::|||:::::::   :::|||:::
     215 Lys.........ArgSerValPheAlaProLeuPheAlaThrGlnLysAl 228

     663 TACCATGACGCTGGCGGCAAAATTGGCACACGTCAAAGGC......GTGA 706
          |||:::   :::   :::||||||||||:::      :::      |||
     228 aThrLeuProValMetGlyLysLeuAlaGluLysThrAsnAlaLeuValV 245

     707 AAACCCTGTTTTTCTGCTGCGAACGCCTGCCTGGCGGACAAGGTTTCGAT 756

         ::::::   :::   ::::::      |||         |||:::
     245 alProValTyrAlaAlaTyrAsnGluSerLeuGly......LysPheGlu 259

     757 TTGCACATCCGCCCCC...GTCCAAGGGGAATTGAACGGCGACAAAGCCCA 803
             |||||||||   :::|||:::      :::   :::      ::
     260 ThrPheIleArgProAlaMetGlnAsnPheProSerGluSerProGluGl 276

     804 TGATGCCGCCGTGTTCAACCGCAATGCCGAATATTGGATACGCCGTTTTC 853
         :|||||:::::::::|||:::::   |||      |||:::
     276 nAspAlaValMetMetAsnLysGluIleGluAlaLeuIleGluCysGlyV 293

     854 CGACGCAGTATCTGTTT 870
             ||||||::::::
     293 alAspGlnTyrMetTrp 298
```